

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-168219

(43)Date of publication of application : 22.06.2001

(51)Int.Cl. H01L 21/8247  
H01L 29/788  
H01L 29/792  
H01L 27/115

(21)Application number : 11-361877 (71)Applicant : SONY CORP

(22)Date of filing : 20.12.1999 (72)Inventor : FUJIWARA ICHIRO

(30)Priority

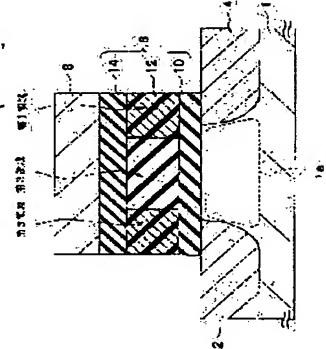
Priority number : 11277642 Priority date : 29.09.1999 Priority country : JP

## (54) NONVOLATILE SEMICONDUCTOR STORAGE DEVICE AND ITS DRIVING METHOD

### (57)Abstract:

PROBLEM TO BE SOLVED: To enhance a write speed, keeping an operation voltage low, in a MONOS-type memory transistor.

SOLUTION: This device is equipped with a substrate 1, a channel formation region 1a for a semiconductor provided at the surface of the substrate, first and second impurity regions 2 and 4 made at the surface of the substrate with the channel formation region 1a in-between, a gate insulating film 6 including within a charge accumulating means (carrier trap) dispersed in the direction of interior of the face opposed to the channel formation area 1a and the direction of film thickness, and a gate electrode 8 provided on the gate insulating film 6. Charge is accelerated in the vertical direction to the substrate such as substrate hot electron, secondary collision ionized hot electron, or the like, or a step 1b is made at the surface of the channel formation region 1a. As a result, the charge accumulating means comes to be positioned in the direction of charge accumulation, and implantation efficiency rises.



### LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

\*.NOTICES \*

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## CLAIMS

---

[Claim(s)]

[Claim 1] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film The non-volatile semiconductor memory which has the charge storage means established in the direction in which it is discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness, and a charge is accelerated at the time of actuation.

[Claim 2] The non-volatile semiconductor memory according to claim 1 whose charge poured into the above-mentioned charge storage means is a hot electron.

[Claim 3] The non-volatile semiconductor memory according to claim 1 to which acceleration of the above-mentioned charge is mainly performed by the perpendicular electric-field component between the 1st impurity range of the above, and the above-mentioned gate electrode.

[Claim 4] The above-mentioned charge storage means is a non-volatile semiconductor memory according to claim 3 which has the 1st field by the side of the 1st impurity range of the above where the above-mentioned charge is poured in at the time of actuation, and the 2nd field where a charge is not poured in.

[Claim 5] The above-mentioned charge storage means is a non-volatile semiconductor memory according to claim 4 which has the 3rd field where a charge is poured into the location by the side of the 2nd impurity range of the above which confronted in the 1st field of the above, and the direction of a channel each other across the 2nd field of the above at the time of actuation.

[Claim 6] The charge poured into the above-mentioned charge storage means is a non-volatile semiconductor memory according to claim 3 which is the hot electron generated by secondary-impact ionization.

[Claim 7] The charge poured into the above-mentioned charge storage means is a non-volatile semiconductor memory according to claim 3 which is the substrate hot electron which was accelerated within the depletion layer of the above-mentioned channel formation field, and was generated.

[Claim 8] The non-volatile semiconductor memory according to claim 1 to which the level difference which makes the 1st impurity range of the above low relatively to the 2nd impurity range of the above is prepared in the front face of the above-mentioned channel formation field, and acceleration of the above-mentioned charge is mainly performed along a channel formation field by the upper part side of the above-mentioned level difference by the level electric-field component between the 1st and 2nd impurity ranges of the above.

[Claim 9] The non-volatile semiconductor memory according to claim 8 with which the above-mentioned charge storage means is formed focusing on the gate-dielectric-film part between the above-mentioned level difference and the 1st impurity range of the above.

[Claim 10] The non-volatile semiconductor memory according to claim 8 whose channel lay length of the

above-mentioned level difference upper part where the above-mentioned charge is accelerated is below the mean free path of the charge in a channel.

[Claim 11] The non-volatile semiconductor memory according to claim 8 set as within the limits below the predetermined distance the channel lay length of the above-mentioned level difference upper part where the above-mentioned charge is accelerated can run, without influencing the charge in a channel of impurity scattering.

[Claim 12] The non-volatile semiconductor memory according to claim 8 whose channel lay length of the above-mentioned level difference upper part where the above-mentioned charge is accelerated is 50nm or less.

[Claim 13] Said charge storage means is a non-volatile semiconductor memory according to claim 1 which does not have the conductivity as the whole field which counters said channel formation field when there is no migration of a charge between the exteriors at least.

[Claim 14] Said gate dielectric film is a non-volatile semiconductor memory containing the bottom insulator layer on said channel formation field, and the nitride or oxidation nitride on the bottom insulator layer concerned according to claim 13.

[Claim 15] Said gate dielectric film is a non-volatile semiconductor memory containing the diameter conductor of a granule which was formed on the bottom insulator layer as said charge storage means, and was insulated with the bottom insulator layer on said channel formation field according to claim 13.

[Claim 16] The non-volatile semiconductor memory according to claim 15 whose particle size of said diameter conductor of a granule is 10 nanometers or less.

[Claim 17] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film The non-volatile semiconductor memory with which it has the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness, and the level difference which makes the 1st impurity range of the above low relatively to the 2nd impurity range of the above is prepared in the front face of the above-mentioned channel formation field.

[Claim 18] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It has the charge storage means discretized in the inside of the field which counters the above-mentioned channel formation field, and the direction of thickness. The above-mentioned gate electrode The 1st gate electrode by which the laminating was carried out to the 1st impurity range side of the above of the above-mentioned channel formation field through the 1st gate dielectric film which does not have a charge storage means inside, The non-volatile semiconductor memory containing the 2nd gate electrode by which the laminating was carried out to the 2nd impurity range side of the above of the above-mentioned channel formation field through the 2nd gate dielectric film with a charge storage means inside.

[Claim 19] The non-volatile semiconductor memory according to claim 18 with which the above-mentioned 2nd gate electrode is formed in the side attachment wall of the above-mentioned 1st gate electrode through the insulator layer.

[Claim 20] The non-volatile semiconductor memory according to claim 18 which equips the front face of the above-mentioned channel formation field with the level difference which makes the 2nd impurity range of the above low relatively to the 1st impurity range of the above.

[Claim 21] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the direction of thickness in opposite \*\*\*\*\* to the above-mentioned channel formation field. The drive approach of the non-volatile semiconductor memory poured in into the charge storage means by which discretization was carried out [ above-mentioned ] where it accelerated the charge within the depletion layer around the above-mentioned channel formation field and momentum is held at the time of actuation.

[Claim 22] The drive approach of a non-volatile semiconductor memory according to claim 21 of injecting into the part by the side of the 1st impurity range of the above of the above-mentioned charge storage means the charge by which acceleration was carried out [ above-mentioned ].

[Claim 23] The drive approach of the non-volatile semiconductor memory according to claim 22 which makes reverse the bias impression conditions over the 1st and 2nd impurity range of the above, and is injected into the part by the side of the 2nd impurity range of the above of the above-mentioned charge storage means.

[Claim 24] The drive approach of a non-volatile semiconductor memory according to claim 21 that the mileage from acceleration initiation of the above-mentioned charge to impregnation is 50nm or less.

[Claim 25] The drive approach of a non-volatile semiconductor memory according to claim 21 of carrying out the electric conduction of the above-mentioned charge to tic [ BARISU ] within a channel.

[Claim 26] The drive approach of the non-volatile semiconductor memory according to claim 21 which controls acceleration and impregnation of the above-mentioned charge by the separate gate electrode.

[Claim 27] The drive approach of the non-volatile semiconductor memory according to claim 21 which makes less than [ 5V ] maximum of the electrical potential difference impressed to the above-mentioned gate electrode.

[Claim 28] The drive approach of the non-volatile semiconductor memory according to claim 27 which makes less than [ 5V ] maximum of the electrical potential difference impressed to the above-mentioned channel formation field, the 1st and 2nd impurity range of the above, and the above-mentioned gate electrode.

[Claim 29] The drive approach of a non-volatile semiconductor memory according to claim 21 of impressing a predetermined read-out drain electrical potential difference between the 1st and 2nd impurity ranges of the above so that the 1st impurity range of the above may serve as the source, and impressing predetermined read-out gate voltage to the above-mentioned gate electrode at the time of read-out.

[Claim 30] The drive approach of the non-volatile semiconductor memory according to claim 23 which changes the electrical-potential-difference impression direction and reads the two or more bits data memorized according to the charge partially poured into the above-mentioned charge storage means between the 1st and 2nd impurity range of the above.

[Claim 31] The drive approach of the non-volatile semiconductor memory according to claim 22 which keeps lengthening the charge which is poured in from the 1st impurity range side of the above at the time of elimination, and is held for the above-mentioned charge storage means at it to the 1st impurity range side by direct tunneling or FN tunneling.

[Claim 32] The drive approach of the non-volatile semiconductor memory according to claim 23 which is poured in from the 1st or 2nd impurity range side of the above at the time of elimination, bundles up individually the charge which is divided into the above-mentioned charge storage means by the both sides of the direction of a channel, and is held by direct tunneling or FN tunneling, and it keeps lengthening to a substrate side.

[Claim 33] Said charge storage means is the drive approach of the non-volatile semiconductor memory according to claim 21 which does not have the conductivity as the whole field which counters said channel formation field when there is no migration of a charge between the exteriors at least.

[Claim 34] Said gate dielectric film is the drive approach of the non-volatile semiconductor memory containing the bottom insulator layer on said channel formation field, and the nitride or oxidation nitride on the bottom insulator layer concerned according to claim 33.

[Claim 35] Said gate dielectric film is the drive approach of the non-volatile semiconductor memory containing the diameter conductor of a granule which was formed on the bottom insulator layer as said charge storage means, and was insulated with the bottom insulator layer on said channel formation field according to claim 33.

[Claim 36] The drive approach of a non-volatile semiconductor memory according to claim 35 that the particle size of said diameter conductor of a granule is 10 nanometers or less.

[Claim 37] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness. a charge is accelerated within the channel made to the above-mentioned channel formation field at the time of actuation -- making -- BARISU -- the drive approach of the non-volatile semiconductor memory poured in into the charge storage means by which discretization was carried out [ above-mentioned ] using the electric conduction phenomenon [ tic ].

[Claim 38] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness. The drive approach of a non-volatile semiconductor memory of pouring in the hot electron generated by secondary-impact ionization at the time of actuation into the charge storage means by which discretization was carried out [ above-mentioned ].

[Claim 39] A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness. The drive approach of the non-volatile semiconductor memory poured in into the charge storage means by which discretization was carried out [ above-mentioned ] by accelerating a charge within the depletion layer made to the above-mentioned channel formation field, and making it a substrate hot electron at the time of actuation.

[Claim 40] A substrate and the 1st and 2nd impurity ranges which are established in the front face of the substrate concerned, are formed in a substrate front face across the channel formation field and the channel formation field concerned of a semi-conductor, and serve as the source or a drain at the time of actuation, The gate dielectric film formed on the above-mentioned channel formation field, and the gate electrode formed on the gate dielectric film concerned, It is discretized in the inside of the field

which countered the above-mentioned channel formation field, and the direction of thickness, and is formed in the above-mentioned gate dielectric film. At the time of actuation, a channel hot electron, a BARISUCHIKKU hot electron, It is the drive approach of a non-volatile semiconductor memory of having a charge storage means by which the hot electron resulting from the tunnel current between a secondary-impact ionization hot electron, a substrate hot electron, or bands is poured in. The charge which is poured in from the above 1st and/or 2nd impurity range side at the time of elimination, and is held on the one side of the direction of a channel, or both sides at the above-mentioned charge storage means at it The drive approach of the non-volatile semiconductor memory which bundles up individually by direct tunneling or FN tunneling, and it keeps lengthening to a substrate side.

---

[Translation done.]

**\* NOTICES \***

**JPO and NCIP are not responsible for any damages caused by the use of this translation.**

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

**[Detailed Description of the Invention]**

**[0001]**

**[Field of the Invention]** This invention inside the gate dielectric film between the channel formation field of a memory transistor, and a gate electrode the charge storage means (for example, the charge trap in the nitride in a MONOS mold or an MNOS mold --) discretized superficially It has a charge trap or a diameter conductor of a granule near the interface of a top insulator layer and a nitride etc., and is related with the non-volatile semiconductor memory which considers pouring in a charge (an electron or hole) electrically, and accumulating it, or drawing it out to the charge storage means concerned as basic actuation, and its drive approach.

**[0002]**

**[Description of the Prior Art]** Although non-volatile semiconductor memory is expected as a small information record medium with large capacity, drawing speed equivalent to a network transmission speed (for example, carrier frequency: 100MHz) is increasingly required with broadband-ization of an information network in recent years. For this reason, improvement in the drawing speed beyond a single figure or it is demanded from non-volatile semiconductor memory from the drawing speed of the conventional 100microsec / cel with good scaling nature.

**[0003]** Non-volatile semiconductor memory is FG (Floating Gate) which a charge storage means to hold a charge followed superficially. MONOS to which the charge storage means other than a mold was discretized superficially (Metal-Oxide-Nitride-Oxide Semiconductor) There is a mold etc.

**[0004]** In MONOS mold non-volatile semiconductor memory, the carrier trap of the interface of the inside of [Nitride Six Ny ( $0 < x < 1, 0 < y < 1$ )] film which is bearing charge maintenance actively or a top insulator layer, and a nitride spatially (in namely, the direction of a field and the direction of thickness) Since it has discretized and spread, a charge maintenance property is Six Ny besides tunnel insulator layer thickness. It is dependent on the distribution [ be / energy-of-a-charge-/ it / spatial ] captured

by the carrier trap in the film.

[0005] Since the charge storage means has been spatially discretized in the MONOS mold to the thing to which many charges leak through leak pass in FG mold, and a charge maintenance property tends to fall when leakage current pass is locally generated in this tunnel insulator layer, it does not pass for the local charge of the leak pass circumference to leak locally through leak pass, but is hard for the charge maintenance property of the whole storage element to fall. For this reason, in a MONOS mold, the problem of a fall of the charge maintenance property by thin-film-izing of a tunnel insulator layer is not so serious as FG mold. Therefore, the scaling nature of the tunnel insulator layer in a detailed memory transistor with very short gate length excels FG mold in the MONOS mold. Moreover, when a charge is locally poured in to the distribution flat surface of the carrier trap discretized superficially, the charge is held, without being spread in the inside of a flat surface, and the direction of thickness like FG mold.

[0006] An improvement of a DISUTABU property is important for realizing a detailed memory cell by MONOS mold nonvolatile memory, and, for that purpose, the tunnel insulator layer will need to be set up more thickly than the usual thickness (1.6nm – 2.0nm).

[0007]

[Problem(s) to be Solved by the Invention] However, in the conventional MONOS mold nonvolatile memory, a tunnel insulator layer is thick-film-ized, or it has become a still more disadvantageous situation for the improvement in drawing speed to reduce operating voltage. For this reason, when dependability (for example, a data-hold property, the Lee Dodi Staab property, or a data rewriting property etc.) is fully satisfied to nonvolatile memory, such as the conventional MONOS mold, 100microsec of drawing speed is a limitation.

[0008] Considering only drawing speed, it is easy to accelerate the impregnation method using a channel hot electron (CHE) from whole channel surface FN tunneling. However, the usual CHE impregnation method which generates CHE at a drain edge is not so enough as injection efficiency as  $1 \times 10^{-6}$ . Moreover, although the source side impregnation mold MONOS transistor which pours in CHE from a source side was reported in order to gather injection efficiency (IEEE Electron Device Letter19, 1998, and pp153), this source side impregnation mold MONOS transistor of dependability, such as 14V, a high top and the Lee Dodi Staab property, and a data rewriting property, is not enough as operating voltage at the time of writing at the time of 12V and elimination.

[0009] Thus, in nonvolatile memory, such as the conventional MONOS mold, it has improvement in drawing speed, and the relation of a trade-off of low-battery-izing and dependability reservation of operating voltage. Conquest of this trade-off had become an important technical problem from a viewpoint of mixed loading with a logical circuit in the system LSI to which development is advanced briskly further in recent years, when developing high-speed mass nonvolatile memory.

[0010] The purposes of this invention are the non-volatile semiconductor memory which has the structure which can be improved in drawing speed, maintaining dependability, such as the Lee Dodi Staab property, good, and low-battery-izing operating voltage, and offering the drive approach in the memory transistor which is made to accumulate a charge in carrier traps discretized superficially, such as a MONOS mold, and carries out basic actuation.

[0011]

[Means for Solving the Problem] The non-volatile semiconductor memory concerning the 1st viewpoint of this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It has the charge storage means established in the direction in which it is discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness, and a charge is accelerated at the time of actuation. Any of a horizontal direction and a perpendicular

direction are sufficient as the direction where this charge is mainly accelerated at a substrate.

[0012] Since the non-volatile semiconductor memory concerning this 1st viewpoint has established the discretized charge storage means in the direction in which a charge is accelerated, the momentum of a charge is easy to be maintained in the case of impregnation. For this reason, a charge is efficiently poured in to a charge storage means.

[0013] The non-volatile semiconductor memory concerning the 2nd viewpoint of this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It has the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness, and the level difference which makes the 1st impurity range of the above low relatively to the 2nd impurity range of the above is prepared in the front face of the above-mentioned channel formation field.

[0014] Suitably, the above-mentioned charge storage means is formed focusing on the gate-dielectric-film part between the above-mentioned level difference and the 1st impurity range of the above. Suitably, the channel lay length (for example, die length from the 2nd impurity range edge of a channel formation field to a level difference) of the above-mentioned level difference upper part where the above-mentioned charge is accelerated is below the mean free path of the electron in a channel. Or this die length is set as 50nm or less within the limits below the predetermined distance it can run, without influencing the electron in a channel of impurity scattering.

[0015] Generally, in the non-volatile semiconductor memory which writes in using hot electron impregnation, when predetermined bias voltage is impressed between the source and a drain and a write-in predetermined electrical potential difference is impressed to a gate electrode, the electric-field acceleration of the charge (electron in this case) supplied from the source in the formed channel is carried out. The charge (hot electron) which obtained high energy near the drain by the acceleration is drawn to the electric field by the gate electrode, and is poured into a charge storage means.

[0016] The non-volatile semiconductor memory concerning the 2nd viewpoint of this invention prepared the level difference in the substrate front face, and has formed the gate dielectric film which has a charge storage means (carrier trap) in the interior, such as ONO film, by the side with a low level difference. Therefore, the high energy charge (for example, hot electron) generated near the drain is efficiently poured into a high speed at a carrier trap, without almost losing kinetic energy, maintaining the momentum (a direction and magnitude). The non-volatile semiconductor memory concerning the 2nd viewpoint shows one gestalt of level charge acceleration of the non-volatile semiconductor memory concerning the 1st viewpoint. If it carries out less than to the channel lay length of the level difference upper part where this charge is accelerated, for example, 50nm, the effect of impurity scattering etc. will be conducted to tic [ BARISU ], without a charge hardly receiving the inside of a channel. Therefore, the effectiveness of charge impregnation and a rate increase further.

[0017] The non-volatile semiconductor memory concerning the 3rd viewpoint of this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It has the charge storage means discretized in the inside of the field which counters the above-mentioned channel formation field, and the direction of thickness. The above-mentioned gate electrode The 1st gate electrode by which the laminating was carried out to the 1st impurity range side of the above of the above-mentioned channel formation field through the 1st gate dielectric film which does not have a

charge storage means inside, The 2nd gate electrode by which the laminating was carried out to the 2nd impurity range side of the above of the above-mentioned channel formation field through the 2nd gate dielectric film with a charge storage means inside is included. Suitably, the 2nd gate electrode is formed in the side attachment wall of the 1st gate electrode through the insulator layer.

[0018] In the non-volatile semiconductor memory concerning this 3rd viewpoint, the 1st gate electrode by the side of the 1st impurity range and the 2nd gate electrode by the side of the 2nd impurity range are prepared separately. Therefore, the gate voltage impression conditions at the time of acceleration and the gate voltage impression conditions at the time of impregnation can be set up separately, and it is easy to carry out control for maximum-izing injection efficiency and accelerating writing so much. Since it is good without forming a charge storage means in the gate dielectric film by the side of the 1st gate electrode (the 1st gate dielectric film), the 1st gate dielectric film is made thin and drift potential can be enlarged. Moreover, fixed, then applied voltage can be lowered for drift potential. On the other hand, when the 2nd gate electrode is made into a sidewall form, cel area is small and ends.

[0019] The drive approach of the non-volatile semiconductor memory concerning the 4th viewpoint of this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the direction of thickness in opposite \*\*\*\*\* to the above-mentioned channel formation field. At the time of actuation, a charge is accelerated within the depletion layer around the above-mentioned channel formation field, and it pours in into the charge storage means by which discretization was carried out [ above-mentioned ] where momentum is held. Suitably, independently partial charge impregnation is performed to the 1st impurity range [ of a charge storage means ], and 2nd impurity range side in the case of writing. Specifically, it writes in by making electrical-potential-difference impression conditions reverse. In read-out, the electrical-potential-difference impression direction between the 1st and 2nd impurity range is decided, and a predetermined read-out drain electrical potential difference is impressed, and predetermined read-out gate voltage is impressed to a gate electrode so that the charge corresponding to the information which should be read may become a source side suitably. In two or more bits read-out, the electrical-potential-difference impression direction between the 1st and 2nd impurity range is made reverse, and such read-out is performed. In elimination, the charge which is poured in from the above 1st and/or 2nd impurity range side, and is held on the one side of the direction of a channel or both sides at the above-mentioned charge storage means is individually put in block by direct tunneling or FN tunneling, and lengthening it to a substrate side is kept.

[0020] The drive approach of the non-volatile semiconductor memory concerning the 5th viewpoint of this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness. a charge is accelerated within the channel made to the above-mentioned channel formation field at the time of actuation -- making -- BARISU -- it pours in into the charge storage means by which discretization was carried out [ above-mentioned ] using the electric conduction phenomenon [ tic ].

[0021] The drive approach of the non-volatile semiconductor memory concerning the 6th viewpoint of

this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness, and the hot electron generated by secondary-impact ionization at the time of actuation is poured in into the charge storage means by which discretization was carried out [ above-mentioned ].

[0022] The drive approach of the non-volatile semiconductor memory concerning the 7th viewpoint of this invention A substrate and the channel formation field of a semi-conductor established in the front face of the substrate concerned, The 1st and 2nd impurity ranges which are formed in a substrate front face across the channel formation field concerned, and serve as the source or a drain at the time of actuation, Within the gate dielectric film formed on the above-mentioned channel formation field, the gate electrode prepared on the gate dielectric film concerned, and the above-mentioned gate dielectric film It is the drive approach of a non-volatile semiconductor memory of having the charge storage means discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness. A charge is accelerated within the depletion layer made to the above-mentioned channel formation field at the time of actuation, and it pours in into the charge storage means by which discretization was carried out [ above-mentioned ] by making it a substrate hot electron.

[0023] The drive approach of the non-volatile semiconductor memory concerning the 8th viewpoint of this invention A substrate and the 1st and 2nd impurity ranges which are established in the front face of the substrate concerned, are formed in a substrate front face across the channel formation field and the channel formation field concerned of a semi-conductor, and serve as the source or a drain at the time of actuation, The gate dielectric film formed on the above-mentioned channel formation field, and the gate electrode formed on the gate dielectric film concerned, It is discretized in the inside of the field which countered the above-mentioned channel formation field, and the direction of thickness, and is formed in the above-mentioned gate dielectric film. At the time of actuation, a channel hot electron, a BARISUCHIKKU hot electron, It is the drive approach of a non-volatile semiconductor memory of having a charge storage means by which the hot electron resulting from the tunnel current between a secondary-impact ionization hot electron, a substrate hot electron, or bands is poured in. The charge which is poured in from the above 1st and/or 2nd impurity range side at the time of elimination, and is held on the one side of the direction of a channel or both sides at the above-mentioned charge storage means at it is individually put in block by direct tunneling or FN tunneling, and lengthening it is kept to a substrate side.

[0024] In order to use the channel hot electron impregnation through a level difference, the charge impregnation using BARISU tic conduction, secondary-impact ionization hot electron impregnation, or substrate hot electron impregnation, for example by the drive approach of the non-volatile semiconductor memory concerning the 4th – the 7th viewpoint, the effectiveness of charge impregnation is good, and even if service voltage is low, writing high-speed enough is possible. In this invention, a charge can also be partially poured in by choosing an impregnation method suitably to the charge storage means which it was horizontal to the substrate, could also carry out the direction where a charge is accelerated perpendicularly, and was discretized superficially. Moreover, by the drive approach of the non-volatile semiconductor memory concerning the 8th viewpoint, since it eliminates by drawing out a charge to a substrate side at tunneling, a hole does not move so much during elimination actuation like before in the inside of a bottom insulator layer. In addition, this invention is suitable for the diameter conductor mold of a granule containing the diameter conductor of a granule each other

insulated on the bottom insulator layer within gate dielectric film, such as a MONOS mold or an MNOS mold which contains a nitride or an oxidization nitride on a bottom insulator layer within gate dielectric film.

[0025]

[Embodiment of the Invention] 1st operation gestalt drawing 1 is the circuit diagram showing the outline configuration of the memory cell array of the non-volatile semiconductor memory of the source line separation NOR mold concerning the operation gestalt of this invention.

[0026] Each memory cell of a NOR mold memory cell array is constituted from one memory transistor by this nonvolatile memory equipment. As shown in drawing 1, the memory transistors M11–M22 are arranged in the shape of a matrix, and between these transistors is wired by the word line, the bit line, and the discrete-type source line. That is, each drain of the memory transistors M11 and M12 which adjoin in the direction of a bit is connected to a bit line BL1, and each source is connected to the source line SL 1. Similarly each drain of the memory transistors M21 and M22 which adjoin in the direction of a bit is connected to a bit line BL2, and each source is connected to the source line SL 2. Moreover, each gate of the memory transistors M11 and M21 which adjoin in the direction of WORD is connected to a word line WL1, and each gate of the memory transistors M12 and M22 which adjoin in the direction of WORD is similarly connected to the word line WL2. In the whole memory cell array, such cel arrangement and connection between cels are repeated.

[0027] Drawing 2 is the outline top view of the detailed NOR mold cel array concerning the 1st operation gestalt. Moreover, drawing 3 is the bird's-eye view seen from the cross-section side which met the A-A' line of drawing 2.

[0028] In this detailed NOR mold memory cell array, as shown in drawing 3, the isolation insulating layer ISO is formed in the front face of the semi-conductor substrate SUB of n mold or p mold (n wells or p wells are also good) from a trench or LOCOS. The isolation insulating layer ISO is arranged in the shape of [ long in the direction of a bit (lengthwise direction of drawing 2 ) ] an parallel stripe, as shown in drawing 2. It intersects perpendicularly with the isolation insulating layer ISO mostly, and each word lines WL1, WL2, WL3, and WL4 and -- are wired at equal intervals. This word line carries out the laminating of the gate dielectric film which consists of a bottom insulator layer, a nitride, and a top insulator layer, and the gate electrode, and is constituted so that it may mention later.

[0029] the active region in spacing of each isolation insulating layer ISO -- setting -- alienation of each word line -- the impurity of a substrate 101 and a reverse conductivity type is introduced into high concentration, and the source impurity range (the 2nd impurity range) S and the drain impurity range (the 1st impurity range) D are formed in the tooth space by turns. That magnitude is specified in the direction of WORD (longitudinal direction of drawing 2 ) only at intervals of the isolation insulating layer ISO, and this source impurity range S and the drain impurity range D are prescribed in the direction of a bit by only WORD line spacing. Therefore, the source impurity range S and the drain impurity range D are extremely formed in homogeneity from the error of mask alignment hardly being introduced about dispersion in the magnitude and arrangement.

[0030] The upper part and the side attachment wall of a word line are covered by the insulating layer. That is, an offset insulating layer is arranged by word lines WL1 and WL2 and the same pattern as the upper part of --, and the sidewall insulating layer is formed in the both-sides wall of the laminating pattern which consists of an offset insulating layer, a gate electrode under it (word line), and gate dielectric film. Along with the word line, opening of the long and slender self-align contact hole is carried out to the tooth-space part of each word lines by this offset insulating layer and the sidewall insulating layer.

[0031] In the self-align contact hole, a conductive ingredient is embedded alternately and, thereby, the bit contact BC and the source contact SC are formed so that it may lap with the source impurity range S or the drain impurity range D in part. In formation of these contacts BC and SC, an electrical conducting material is deposited so that the inside of a self-align contact hole may be embedded, and

the resist pattern for etching masks is formed on it. At this time, a resist pattern is made somewhat larger than the width of face of a self-align contact hole, and a part is put on the isolation insulating layer ISO. And etching removes the electrical conducting material of the perimeter of a resist pattern by using this resist pattern as a mask. Thereby, two kinds of contacts BC and SC are formed in coincidence.

[0032] The crevice of the perimeter of contact is embedded by the insulator layer which is not illustrated. In this insulator layer top, the bit lines BL1 and BL2 which contact on the bit contact BC, --, the source lines SL1 and SL2 and -- which contact on the source contact SC are formed in the shape of an parallel stripe by turns.

[0033] Contact formation of as opposed to that bit line or a source line in this detailed NOR mold cel array is attained by formation of a self-align contact hole, and formation of a plug. In formation of a self-align contact hole, while insulating separation with a word line is attained, the expressional side of the source impurity range S or the drain impurity range D is formed in homogeneity. And formation of the bit contact BC and the source contact SC is performed to the expressional side of the source impurity range S in this self-align contact hole, or the drain impurity range D. Therefore, the size of the direction of a bit is mostly decided by formation of a self-align contact hole, and, as for the substrate contact surface of each plug, the variation in the part and contact area is small.

[0034] Insulating separation with the bit contact BC or the source contact SC, and a word line is easy. That is, it bundles up at the time of word line formation, the offset insulating layer is formed, and a sidewall insulating layer is formed membrane formation of an insulator layer, and only by performing overall etching (etchback) after that. Moreover, further, the bit contact BC, the source contact SC, and since a bit line and a source line carry out pattern NINGU of the same hierarchy's conductive layer and are formed, wiring structure is very simple, and a routing counter also has it, and it is structure advantageous to holding down a manufacturing cost low. [ little ] And since there was almost no useless space, when each class is formed with the minimum line width F of a wafer process limitation, it is 8F2. It can manufacture in a near, very small cel area.

[0035] Drawing 4 is the sectional view showing the component structure of the MONOS mold memory transistor concerning this operation gestalt.

[0036] In a sign 1, the semi-conductor substrates SUB, such as a silicon wafer which has the conductivity type of n mold or p mold, or a well (henceforth a substrate), and 1a show a channel formation field among drawing 4 , and 2 and 4 show the source impurity range S of the memory transistor concerned, and the drain impurity range D. A "channel formation field" means the field where the channel to which an electron or an electron hole conducts current is formed in the interior of a front-face side by this invention. The part by which the "channel formation field" of this example was inserted into the source impurity range 2 and the drain impurity range 4 within the substrate corresponds.

[0037] The conductivity formed by introducing the impurity of channel formation field 1a and a reverse conductivity type into a substrate 1 at high concentration is a high field, and the source impurity range 2 and the drain impurity range 4 have various gestalten. Usually, it is LDD (Lightly Doped Drain) to the substrate surface location which attends channel formation field 1a of the source impurity range 2 and the drain impurity range 4. The low concentration field to call is made to provide in many cases.

[0038] On channel formation field 1a, the laminating of the gate electrode 8 of a memory transistor is carried out through gate dielectric film 6. Generally the gate electrode 8 is the polish recon (doped poly-Si) by which the impurity of p mold or n mold was introduced into high concentration, and was electric-conduction-ized, or doped poly-Si. It consists of a cascade screen with refractory metal silicide. The channel lay length (gate length) of this gate electrode 8 is about 0.25 micrometers or less, for example, 0.18 micrometers.

[0039] The gate dielectric film 6 in this operation gestalt consists of a bottom insulator layer 10, a nitride 12, and a top insulator layer 14 sequentially from the lower layer. The bottom insulator layer 10

forms an oxide film, it carries out nitriding treatment of this, and it is used for it. The thickness of the bottom insulator layer 10 can be decided within the limits of 2.0 to 5.0nm according to a use application, and is set as 5.0nm here.

[0040] The nitride 12 consists of 5.0nm silicon nitride (Six Ny) film ( $0 < x < 1$ ,  $0 < y < 1$ ). It is produced by reduced pressure CVD (LP-CVD), many carrier traps are contained in the film, and this nitride 12 shows the electric conduction property of a pool FURENKERU mold (PF mold).

[0041] The top insulator layer 14 needs to form a deep carrier trap in high density near the interface with a nitride 12, for this reason, oxidizes the nitride after membrane formation thermally, and is formed. moreover, the top insulator layer 14 -- HTO (High Temperature chemical vapor deposited Oxide) -- SiO<sub>2</sub> formed by law It is good also as film. When the top insulator layer 14 is formed by CVD, this trap is formed of heat treatment. In order to prevent impregnation of the hole from the gate electrode 8 effectively and to aim at fall prevention of the count in which data rewriting is possible, also at the lowest, 3.0nm 3.5nm or more is preferably required for the thickness of the top insulator layer 14. Here, top insulator layer thickness is set to 3.5nm.

[0042] the semi-conductor substrate first prepared in manufacture of the memory transistor of such a configuration -- the isolation insulating layer ISO and a well -- after forming W and performing the ion implantation for threshold voltage adjustment etc. if needed, gate dielectric film 6 is formed. Concretely, heat treatment of 1000 degrees C and 10sec is performed by the short-time elevated-temperature heat-treating method (RTO law), and the silicon oxide film (bottom insulator layer 10) is formed. Next, on the bottom insulator layer 10, with LP-CVD method, a silicon nitride film (nitride 12) is deposited more thickly than this so that the last thickness may be set to 5nm. This CVD is performed at the substrate temperature of 650 degrees C using the gas which mixed dichlorosilane (DCS) and ammonia. The formed silicon nitride film front face is oxidized by the oxidizing [ thermally ] method, for example, the 3.5nm silicon oxide film (top insulator layer 14) is formed. This thermal oxidation is performed at 950 degrees C whenever [ coke-oven-temperature ] for example, into an H<sub>2</sub> O ambient atmosphere. Thereby, the deep carrier trap whose trap level (energy difference from the conduction band of a silicon nitride film) is extent of 2.0eV or less is about one to  $2 \times 10^{13} / \text{cm}^2$ . It is formed by the consistency. Moreover, about 1.6nm (top insulator layer 14) of thermal oxidation silicon film is formed for a nitride 12 to 1nm, the nitride thickness of a substrate decreases at this rate, and the last thickness of a nitride 12 is set to 5nm.

[0043] The laminating of the electric conduction film used as the gate electrode 8 and the cascade screen of an offset insulating layer (un-illustrating) is carried out. And this gate dielectric film 6, the electric conduction film, and the cascade screen of an offset insulating layer are collectively processed by the same pattern. The source and the drain impurity ranges 2 and 4 are formed with ion-implantation in self align with the formed laminating pattern.

[0044] Then, in order to consider as the memory cell array structure of drawing 3 , a self-align contact hole is formed with a sidewall insulating layer, and the bit contact BC and the source contact SC are formed on the source expressed by the self-align contact hole, and the drain impurity ranges S and D (the source and drain impurity ranges 2 and 4). Then, after embedding the perimeter of these contacts with an interlayer insulation film and forming a bit line BL1, -- and the source line SL 1, and -- on an interlayer insulation film, the non-volatile memory cell array concerned is completed through formation and overcoat membrane formation, a pad opening process, etc. through the layer insulation layer performed if needed of the upper wiring.

[0045] Below, the case where data are written in the memory transistor M11 is explained to an example about the example of a bias setting at the time of the writing of the nonvolatile memory of such a configuration, and actuation. With this operation gestalt, writing is performed using the hot electron generated by secondary-impact ionization. In secondary-impact ionization hot electron impregnation, when the electron hole used as the substrate current generated near the drain crosses the depletion layer near the drain and is injected into a substrate, by electric field, it is mainly accelerated

perpendicularly and the electron which received energy from electric field within the depletion layer, and was made to generate an electron and an electron hole pair, among these was generated is poured into a charge storage means.

[0046] At the time of writing, in drawing 4 , 5V are impressed to 3.3V and the gate electrode 8, and 3V are impressed to a well 1 to 0V and the drain impurity range 4 at the source impurity range 2. The electron hole which was poured in from the drain impurity range 4 and went into the nearby depletion layer collides with a silicon atom under this condition, and an electron and an electron hole pair occur within the depletion layer concerned by this collision. Among those, although dissipation of the electron hole is carried out to a substrate side with more low potential, it is accelerated by the electric field between a gate electrode and a substrate, and an electron is accelerated toward the upper part in a depletion layer. An electron serves as a hot electron, and that part overcomes the energy barrier of the bottom insulator layer 10, and it is poured into the charge storage means in a nitride 12 (carrier trap) by this acceleration. At this time, distribution of an impregnation electron carries out localization to the 1st field in the side  $R > 4$  into which the electron hole which causes a collision is poured, i.e., drawing 4 . For this reason, it writes in, when the threshold electrical potential difference of the memory transistor M11 concerned is in an elimination condition, and it goes up to a condition, and writing is made. Page writing is possible by writing in with this writing, changing prohibition and setting up electrical-potential-difference impression conditions for it suitably to two or more memory cells which stand in a row in a selection word line.

[0047] It changes into extent in which a bias value is written and a channel is formed according to a condition in read-out. For example, where the drain impurity range 4 is grounded, 1.5V are impressed to the source impurity range 2, and 2V are impressed to the gate electrode 8. Thereby, in page read-out, a channel is formed in the 1st field of a charge storage means with the memory transistor in the elimination condition that the electron is not poured in, and a channel is not formed with the memory transistor in the write-in condition that the electron was poured into the 1st field of a charge storage means. Therefore, to a bit line BL1 and --, the potential change according to the existence of the channel formation appears. If a sense amplifier detects this potential change, it will be collectively read by the stored data in a page.

[0048] In elimination, it carries out by keeping lengthening a charge using FN tunneling or direct tunneling from the whole channel surface or the drain impurity range 4 side.

[0049] On the other hand, in order to memorize 2-bit data in 1 cel, when performing the same writing also as the source impurity range side of a charge storage means, in this 2nd writing, applied voltage of the source and a drain is made contrary to the 1st time. The same secondary-impact ionization as the 1st time takes place by the source impurity range side by this, and the generated electron is poured into the field by the side of the source impurity range of a charge storage means (the 3rd field). In the cel of a write-in condition, a hot electron is poured into the 1st field of a charge storage means, and it is held, and a hot electron is poured into the 3rd field independently of this, and 2 bits is held. That is, since the 2nd field where a hot electron is not poured in intervenes between the 1st field of a charge storage means, and the 3rd field, the hot electron corresponding to this 2-bit information is distinguished sharply certainly.

[0050] With the case of the charge of the 1st field, read-out of the charge held to the 3rd field makes reverse the electrical-potential-difference direction between the source and a drain, and performs it. Thereby, 2-bit data can be read independently. Elimination is also made reverse and performs elimination by the side of the 1st above mentioned field, and applied voltage of the source impurity range 2 and the drain impurity range 4. In addition, when [ both ] eliminating all over a channel, it is making the source and the drain impurity ranges 2 and 4 into a substrate and this potential, and package elimination of the data by the side of the 1st field and the 3rd field is carried out.

[0051] The current-voltage characteristic of the memory transistor of a write-in condition and an elimination condition was examined. Consequently, the off leak current value from the non-choosing cel

of drain electrical-potential-difference 1.5V was about 1 nA. Since the read-out current in this case is more than 10microA, incorrect read-out of a non-choosing cel does not produce it. Therefore, it read in the MONOS mold memory transistor of 0.18 micrometers of gate length, and the margin of the punch-through pressure-proofing at the time found a certain thing enough. Moreover, the Lee Dodi Staab property of gate voltage 1.5V was evaluated, and it turned out also after the time amount progress more than 3x108 sec that read-out is possible.

[0052] Since the carrier trap is discretized spatially, it is good, and the count of data rewriting is 1x106. It turned out that a time is satisfied. Moreover, a data-hold property is 1x106. It is after data rewriting of a time and 85 degrees C and ten years were satisfied.

[0053] As mentioned above, it was able to confirm that property sufficient as 0.18 micrometer gate length's MONOS mold nonvolatile memory transistor was acquired.

[0054] Impregnation of a hot electron is possible with a drain current low in comparison to the charge storage means discretized by the secondary-impact ionization hot electron pouring-in method in this operation gestalt. Therefore, the punch-through at the time of writing is controlled as compared with the usual hot electron impregnation, consequently there is an advantage that gate length's scaling is easy. Moreover, since impregnation is performed while the momentum of an acceleration charge had been maintained, since a charge was mainly accelerated to a substrate and a perpendicular direction, compared with the usual CHE impregnation method, the injection efficiency of a charge is high.

[0055] The 2nd operation gestalt of the 2nd operation gestalt is related with the nonvolatile memory equipment of an imaginary earth NOR mold. When an imaginary earth mold is divided roughly, it has a split-gate mold and an AMG (Alternate Metal Virtual Ground) mold. A split-gate mold writes in the function of a selection transistor by making it provide in a memory transistor substantially, and prevents DISUTABU. Write-in DISUTABU of a contiguity cel is prevented by an AMG mold connecting diffusion layer wiring which consists of a semi-conductor impurity range to metal wiring every other, for example, making metal wiring a bit line, and choosing diffusion layer wiring during metal wiring as a source line.

[0056] Drawing 5 is the circuit diagram showing the memory cell array configuration of an AMG mold among imaginary earth NOR molds. In this memory cell array, the source line is communalized between contiguity memory cells. This communalized source line functions as a bit line, when operating the next memory cell. Therefore, in this memory cell array, all wiring of the direction of a bit is called a "bit line." Each bit lines BL1-BL3 are diffusion layer wiring which consists of an impurity range of a semi-conductor. Every other [ the / BL1 and BL3 ], for example, bit lines, is connected to the upper metal wiring through the bit contact which is not illustrated.

[0057] Thus, except that cel array structures differ, the base of MONOS mold memory transistor structure and writing, read-out, and elimination actuation is the same as the 1st operation gestalt. Since a source line is substituted for and used with the bit line of the adjoining cel in the case of an imaginary earth NOR mold, the size of the direction of a bit can be reduced compared with a separation source line type, and it is cel area 6F2. It can attain. Being able to record 2-bit data on one memory cell physically, the cel area per bit is 3F2 effectually in this case. It becomes.

[0058] With the 3rd operation gestalt book operation gestalt, whole channel surface tunnel erase is performed with channel hot electron (CHE) impregnation writing. Writing is the usual CHE impregnation and a detail here is omitted.

[0059] A memory transistor is considered as the same configuration as drawing 4 , when drawing out the electron held to the 1st field of a charge storage means using direct tunneling from the whole channel surface, 8V and the source impurity range 2 are impressed to 0V and the drain impurity range 4, and the electrical potential difference of 8V is impressed to the gate electrode 8 at opening and a well 1. Cel elimination is performed by the electron currently held to the 1st field of a charge storage means being drawn out by this at a substrate side. At this time, erasing speed was 1msec extent. Moreover, when eliminating the charge of the 3rd field, elimination by the side of the 1st field and applied voltage of the source impurity range 2 and the drain impurity range 4 are made reverse, and are performed.

Furthermore, when the charge is held to the 1st and 3rd field, in case it eliminates all over a channel, package elimination of the data by the side of the 1st field and the 3rd field may be carried out by making both the source and the drain impurity ranges 2 and 4 into a substrate and this potential. [0060] by this elimination approach, it is markedly alike, and there are few amounts of holes which pass the bottom insulator layer 10 working compared with hot hole impregnation etc., therefore the bottom insulator layer 10 cannot deteriorate easily, and it excels in dependability and endurance (for example, ENDI lance property).

[0061] The 4th operation gestalt of the 4th operation gestalt explains a MONOS mold memory transistor as an example about the method write-in [ high-speed ] of having used the substrate hot electron impregnation phenomenon.

[0062] By the substrate hot electron impregnation method, the source and a drain are impressed with this potential, substrate bias voltage is impressed, and where a thick depletion layer is formed in the bottom of a gate electrode, a charge (electron) is poured into the depletion-ized channel formation field. It is accelerated within a depletion layer, and the poured-in electron obtains the energy more than the obstruction energy of an insulator layer, and is poured into the charge storage means superficially discretized exceeding the insulator layer, and, thereby, writing is made.

[0063] The 1st configuration of the MONOS mold memory transistor concerning this operation gestalt is the same as drawing 4 in the 1st operation gestalt. With this operation gestalt, although the basic configuration is the same as the 1st operation gestalt, each configuration film 10, 12, and 14 of gate dielectric film 6 here is set to 3/5/3.5nm. Moreover, production of this gate dielectric film 6 and production of a memory cell are performed using the same equipment as the above mentioned operation gestalt, and process conditions.

[0064] Below, memory cell actuation is explained. AC bias voltage by which bias of the pn junction concerned is carried out by turns to the pn junction between the source impurity range 2 and a well 1 in the forward direction and hard flow is impressed. In this case, when pn junction is forward bias, an electron is poured into p wells 1 from pn junction. a well -- the inner channel formation field is made to have depletion-ized by substrate bias For this reason, it is accelerated with the positive bias electrical potential difference impressed to the gate, hot carrier impregnation of the electron poured in from pn junction is carried out at the carrier trap in the nitride 12 which is a charge storage means exceeding the obstruction potential of the bottom insulator layer 10, and, thereby, writing is made.

[0065] for example, a drain -- opening and a well -- where potential is held by 0V, as positive bias electrical-potential-difference 5V and AC bias voltage, to the source impurity range 2, it impressed in the forward direction by 0.7V, and the pulse voltage was impressed to the gate electrode 8 by VCC to hard flow. In that case, hot carrier impregnation was performed at the high speed, it wrote in less than [ operating voltage 5V ], and 1micro below of time amount sec has been realized. Moreover, about dependability, the property equivalent to the MONOS mold memory cell of the conventional FN tunnel impregnation method was acquired.

[0066] Since the charge storage means is discretized superficially, the partial impregnation to a nitride 12 other than the impregnation to the nitride 12 whole surface is possible by changing the impregnation conditions of a substrate hot electron. For example, although the charge impregnation to a part for the source flank of a charge storage means (the 1st field) is made by the impregnation from a source side as mentioned above, charge impregnation is possible to a part for the drain flank of a charge storage means (the 3rd field) by pouring in a charge from a drain side as source opening conversely. Since the 2nd field where a charge is not poured in is formed between the 1st field and the 2nd field in the case of drawing 4 , sharp distinction of the information which is 2 bits is possible. Moreover, since the threshold electrical potential difference of the 2nd field by which drawing of the charge in the case of elimination is also performed locally, and charge impregnation is not carried out does not change, superfluous elimination is prevented and it has the advantage that the convergency of the threshold electrical potential difference of an elimination condition is high. On the other hand, to the source and a drain, by

[the] impressing AC bias voltage with the same phase with a big electrical-potential-difference value, the 1st and 3rd field of drawing 4 can be connected, and a charge can be poured into the whole charge storage means.

[0067] The 2nd example of structure of the MONOS mold memory transistor which starts this operation gestalt at drawing 6 is shown. the memory transistor shown in drawing 6 -- a duplex -- a well -- it has structure. namely, the p type semiconductor substrate 1 -- n -- a well 60 forms -- having -- n -- the inside of a well 60 -- p -- a well 61 forms -- having -- this p -- the memory transistor is formed in the well 61. Other basic configurations are the same as drawing 4. However, the thickness specification of each configuration film 10, 12, and 14 of the gate dielectric film 6 of this example was set to 3.5/5/4nm.

[0068] n -- a well 60 and p -- an electron is poured into the channel formation field depletion-ized by carrying out forward bias of the pn junction which consists of wells 61, and this electron is injected into the carrier trap in a nitride 12 by substrate hot electron impregnation, after carrying out electric-field acceleration at a gate electrode side.

[0069] for example, the condition of having considered as drain opening -- gate voltage 5V and the source electrical potential difference VCC -- setting up -- n -- a well 60 and p -- 0.7V were impressed to the forward direction between wells 61. Thereby, it wrote in less than [ operating voltage 5V ], and 1micro below of time amount sec has been realized. Moreover, about the dependability of a memory cell, the property equivalent to the MONOS mold memory cell of the conventional FN tunnel impregnation method was acquired.

[0070] The top where operating voltage is as low as less than [ 5V ], since acceleration is performed mainly at right angles to a substrate, substrate hot electron impregnation as well as the case of secondary-impact ionization has the advantage that injection efficiency is high.

[0071] With the 5th operation gestalt of the 5th operation gestalt, a level difference is prepared in a channel formation field, and CHE impregnation is performed from this level difference. The sectional view of the MONOS memory transistor which starts the 5th operation gestalt at drawing 7 is shown.

[0072] A different point from the memory transistor which requires this memory transistor for the 1st operation gestalt shown in drawing 4 is that level difference 1b is prepared in the substrate front face of a channel formation field 1a part. Level difference 1b has height of about 5-50nm, for example, 10nm, its source impurity range 2 side is relatively high, and it is formed so that the drain impurity range 4 side may become low. Other configurations are the same as that of the 1st operation gestalt. However, the thickness specification of each configuration film 10, 12, and 14 of gate dielectric film 6 is set to 3.5/8.0/3.5nm here.

[0073] In manufacture of the memory transistor of such a configuration, although a different point from the 1st operation gestalt is forming level difference 1b, the detail is given with the below-mentioned operation gestalt.

[0074] Below, the case where data are written in the memory transistor M11 of drawing 1 is explained to an example about the CHE impregnation write-in actuation to the nonvolatile memory of such a configuration.

[0075] After writing in if needed at the time of writing and setting up an inhibit electrical potential difference, a program electrical potential difference is impressed. For example, 5V are impressed to the subdevice-bit line BL1, and a predetermined electrical potential difference, 3V [ for example, ], is impressed to the non-choosing word line WL2, the non-choosing source line SL 2, and the non-subdevice-bit line BL2 at the time of substrate potential 0V. Moreover, the selection source line SL 1 is held by touch-down potential 0V. In this condition, a predetermined electrical potential difference, 7V [ for example, ], is impressed to the selection word line WL1 at the time of substrate potential 0V.

[0076] An inversion layer (channel) is formed in the channel formation field front face between the source impurity range 2 and the drain impurity range 4 under this write-in condition, and the electric-field acceleration of the electron is poured in and carried out from the source impurity range 2 side into a channel. The accelerated electron serves as a hot electron near a level channel edge, high energy

impregnation of the part is carried out over the potential barrier of the bottom insulator layer 10, and it is captured by the carrier trap in gate dielectric film 6. For this reason, the threshold electrical potential difference of the memory transistor M11 concerned writes in from an elimination condition, it goes up to a condition, and writing is made.

[0077] With the transistor structure concerning this operation gestalt, level difference 1b is prepared in the middle of channel formation field 1a, and gate dielectric film 6 extends in the pars-basilaris-ossis-occipitalis side of level difference 1b. Therefore, if the height of level difference 1b is optimized according to the thickness specification of gate dielectric film 6, the distribution core of a carrier trap can be located on the locus bent by the electric field by the production top of the electronic transit direction, or the applied voltage of the gate electrode 8. That is, in the writing of the memory transistor concerning this operation gestalt, the direct impregnation to the carrier trap in the condition of having maintained mostly the momentum (magnitude and direction) of the electron by which electric-field acceleration was carried out is attained. Therefore, the energy loss at the time of impregnation decreases conventionally, efficient charge impregnation is possible at high speed, and improvement in the speed of drawing speed is attained as the result.

[0078] Since the charge which jumped out of the channel by dispersion after being excited in energy in the conventional channel hot electron impregnation was drawn near to a carrier trap side by electric field, the injection efficiency could not but be  $1 \times 10^{-6}$ , and could not but be as low to 1 million pieces a rate as one piece. On the other hand, with the transistor structure concerning this operation gestalt, in channel hot electron impregnation, single or more figures injection efficiency has been improved, and the drawing speed below 10microsec has been attained.

[0079] Furthermore, if distance from the channel length of a part by whom an electron is accelerated, i.e., the edge of the source impurity range 2, to level difference 1b is made short to 50nm or less, since it becomes less than [ it ], this distance's being comparable as the mean free path of a channel electron or the electron by which electric-field acceleration was carried out will carry out the electric conduction of the inside of a channel to tic [ BARISU ]. Since this BARISU tic electron is poured into a high speed in ballistic trajectory at a carrier trap, without being influenced of impurity scattering etc., injection efficiency is made further highly and data writing is accelerated.

[0080] When considering drawing speed as present condition maintenance, it can come out enough by a certain amount of improvement in the speed, and, in a certain case, a channel formation field (a substrate or well) and gate inter-electrode applied voltage can be made into less than [ 10V ] by application of such transistor structure and the approach of writing in. When dividing electrical-potential-difference impression with a gate electrode, a substrate, or a well by the positive supply and the negative supply and performing it, it becomes possible to reduce operating voltage less than [ 5V ] in an absolute value.

[0081] In order that the injection efficiency of a charge may increase, room to reduce the channel current at the time of writing is produced. Therefore, in the conventional channel hot electron impregnation, the juxtaposition writing collectively performed to the memory cell of a large number which were difficult, for example, were connected to the same word line by the limitation of the current drive capacity of a high-voltage circuit was attained.

[0082] In addition, when it carried out about elimination by carrying out a block package by the charge drawing which used FN tunneling from the whole channel surface like usual, the erasing speed was 100msec extent.

[0083] 6th operation gestalt drawing 8 is the sectional view of the MONOS mold memory transistor concerning the 6th operation gestalt.

[0084] The point that this memory transistor differs from the MONOS mold memory transistor of the 5th operation gestalt greatly is that the gate electrode is divided into 1st gate electrode 8a by the side of the source, and 2nd gate electrode 8b by the side of a drain. 1st gate electrode 8a is countered and prepared in the part in which an electron is accelerated, i.e., the level difference 1b upper part, and 2nd

gate electrode 8b is mainly counteracted and prepared in the level difference 1b pars basilaris ossis occipitalis. 1st gate electrode 8a controls generation of the channel which mainly accelerates a charge at the time of writing, and 2nd gate electrode 8b controls the electric field which mainly pour in a charge. The nitride 12 and the top insulator layer 14 have not extended in the 1st gate electrode 8a side among the film which constitutes gate dielectric film 6. That is, only the bottom insulator layer 10 intervenes between 1st gate electrode 8a and channel formation field 1a. On the other hand, between 2nd gate electrode 8b and channel formation field 1a, the gate dielectric film 6 of a three-tiered structure intervenes like the 5th operation gestalt. The clearance between the 1st and 2nd gate electrodes 8a and 8b is embedded by insulating-layer 9a, and sidewall insulating-layer 9b is formed in the lateral surface of the 1st and 2nd gate electrodes 8a and 8b, respectively.

[0085] Of the ion implantation before and behind formation of this sidewall insulating-layer 9b, the impurity range of the LDD structure which consists of source drain impurity ranges 2 and 4 and LDD fields 2a and 4a is formed in the front face of a substrate 1. Moreover, it applies to level difference 1b from the LDD field 4a edge by the side of a drain, and the impurity range 3 of p mold is thinly formed in the front face of channel formation field 1a. in addition, this p type of an impurity range 3 and the above-mentioned LDD fields 2a and 4a — an indispensable configuration — 7 — it is — \*\*

[0086] A sectional view shows an example of the manufacture approach of this MONOS mold memory transistor to drawing 9 (A) – drawing 10 (E).

[0087] After forming an isolation insulating layer, a well, etc. in the front face of a semi-conductor substrate, level difference 1b is formed in the front face of a substrate 1 in drawing 9 (A). In formation of this level difference 1b, like illustration, after covering a part of substrate front face by the mask layer R1, for example, a resist, predetermined depth etching of the silicon front face which is not protected by the resist R1 by dry etching is carried out. Then, the ion implantation which used the same resist R1 as the mask layer is performed, and p mold impurity range 3 is formed in the pars basilaris ossis occipitalis and flank of level difference 1b which were formed.

[0088] By drawing 9 (B), sequential membrane formation of the bottom insulator layer 10, a nitride 12, and the top insulator layer 14 is carried out by the same approach as the 1st operation gestalt after removing a resist R1. Then, after forming the pattern of the wrap mask layer R2, for example, a resist, for the part by the side of the pars basilaris ossis occipitalis, dry etching removes the top insulator layer 14 and nitride 12 of resist R2 perimeter from level difference 1b.

[0089] The electric conduction film which serves as a gate electrode by drawing 9 (C) after removing a resist R2, for example, doped poly-Si, Film 8c is deposited and the wrap mask layer R3, for example, a resist, is formed for the predetermined part of the level difference upper part from on the top insulator layer 14 which carried out pattern NINGU at the previous process, and a nitride 12. It is doped poly-Si of resist R3 perimeter, using this resist R3 as a mask. Film 8c is removed. Then, the ion implantation which used the same resist R3 as the mask layer is performed, and it is doped poly-Si. The low-concentration n mold LDD fields 2a and 4a are formed in the front face of the substrate 1 outside film 8c. In front of this ion implantation or in the back, etching which used the resist R3 as the mask layer removes the bottom insulator layer 10 of a periphery.

[0090] After removal of a resist R3, and at drawing 10 (D), it is doped poly-Si. The resist R4 which has opening which crosses the central part of film 8c is formed. doped poly-Si expressed from the opening by using a resist R4 as a mask. The central part of film 8c is removed. Thereby, 1st gate electrode 8a and 2nd gate electrode 8b dissociate, and are formed.

[0091] After removal of a resist R4, by drawing 10 (E), the insulator layer of for example, a silicon oxide system is deposited on the whole surface, and the perimeter of the 1st and 2nd gate electrodes 8a and 8b is covered on it with an insulating material, and the clearance between two electrodes is embedded with an insulating material on it. By performing whole surface anisotropic etching (etchback) in this condition, 1st and 2nd gate electrode 8a, insulating-layer 9a between 8b, and sidewall insulating-layer 9b are formed in coincidence.

[0092] As shown in drawing 8, the ion implantation of the n mold impurity is carried out to high concentration by using the 1st and 2nd gate electrodes 8a and 8b and insulating layers 9a and 9b as a self-align mask, and the source and the drain impurity ranges 2 and 4 are formed. After that, a memory cell array is completed through many of the same processes as the 1st operation gestalt.

[0093] Writing is made by setting up an electrical potential difference almost like the 5th operation gestalt to the memory transistor of such a configuration. Since it divides into 1st gate electrode 8a for channel formation, and 2nd gate electrode 8b for impregnation electric-field control with this operation gestalt and the gate electrode is prepared at this time, a program electrical potential difference can be set up separately. Therefore, there is an advantage which a bias setup at the time of writing tends to optimize.

[0094] For example, the program electrical potential difference which impresses the program electrical potential difference impressed to 1st gate electrode 8a at the time of writing to 3V and 2nd gate electrode 8b is set as 5V. Thus, after optimizing the bias voltage at the time of writing, when it poured in directly by BARISU tic electronic conduction, it became possible to accelerate drawing speed below 1microsec till about 100ns.

[0095] 7th operation gestalt drawing 11 is the sectional view of the MONOS mold memory transistor concerning the 7th operation gestalt.

[0096] The point that this memory transistor differs from the MONOS mold memory transistor of the 6th operation gestalt greatly is having made the 8d of the 2nd gate electrodes for controlling charge impregnation electric field into the sidewall form. This became possible to make occupancy area of a transistor quite smaller than the case of the 2nd operation gestalt. The nitride 12 and the top insulator layer 14 which constitute gate dielectric film 6 in the insulating layer for 1st and 2nd gate electrode 8a and 8d are used with formation of the 8d of the 2nd gate electrodes of a sidewall form. Other configurations are fundamentally [ as the 6th operation gestalt ] the same.

[0097] In addition, by having made the 8d of the 2nd gate electrodes into the sidewall form, from the case where the distance of the drain impurity range 4 and level difference 1b is the 2nd operation gestalt, it became near and formation of the part and a channel became easy. Therefore, in drawing 11  $R > 1$ , although p mold impurity range is not formed in the front face of channel formation field 1a, of course, p mold impurity range 3 may be formed like the 6th operation gestalt. Moreover, a LDD field may be prepared inside the source and the drain impurity ranges 2 and 4 like the 6th operation gestalt, respectively.

[0098] A sectional view shows an example of the manufacture approach of this MONOS mold memory transistor to drawing 12 (A) – drawing 13 (D).

[0099] First, as shown in drawing 12 (A), level difference 1b is formed in a substrate front face by the same approach as the 6th operation gestalt. Next, the electric conduction film used as the bottom insulator layer 10 and a gate electrode is formed, and pattern NINGU of the electric conduction film used as a gate electrode is carried out by etching which used as the mask the resist which is not illustrated. Thereby, 1st gate electrode 8a is formed in the predetermined location of the upper part of level difference 1b.

[0100] By drawing 12 (B), sequential membrane formation of a nitride 12 and the top insulator layer 14 is carried out by the same approach as the 1st operation gestalt after removing a resist.

[0101] Then, in drawing 13 (C), the electric conduction film used as a gate electrode is thickly deposited on the whole surface, and etchback of this is carried out. Thereby, the conductive layers 8d and 8e of a sidewall form are formed in the both-sides wall of 1st gate electrode 8a through a nitride 12 and the top insulator layer 14.

[0102] In drawing 13 (D), sequential removal of the top insulator layer 14, the nitride 12, and the bottom insulator layer 10 expressing that perimeter is first carried out by using these conductive layers 8d and 8e as a self-align mask. A 8d [ of conductive layers by the side of a level difference pars basilaris ossis occipitalis ] (2nd gate electrode) and 1st gate electrode 8a top is covered, and the mask layer R5, for

example, a resist, is formed. A resist R5 is etched into a mask, and conductive layer 8e of one of the two is removed, then sequential removal of the top insulator layer 14, a nitride 12, and the bottom insulator layer 10 is carried out.

[0103] As shown in drawing 11 after removing a resist R5, the ion implantation of the n mold impurity is carried out to high concentration by using the 1st and 2nd gate electrodes 8a and 8d and insulator layers 12 and 14 as a self-align mask, and the source and the drain impurity ranges 2 and 4 are formed. After that, a memory cell array is completed through many of the same processes as the 1st operation gestalt.

[0104] In the writing to the memory transistor of such a configuration, in spite of having prepared separately the gate electrode the object for channel formation, and for impregnation electric-field control, the occupancy area of the case of the 5th operation gestalt of drawing 7 and a transistor hardly changes. Therefore, there is an advantage that the detailed memory cell suitable for high integration is realizable.

[0105] The electrical-potential-difference setting approach at the time of writing has the 6th operation gestalt and the high field strength of the longitudinal direction corresponding to the applied voltage of 1st gate electrode 8a although it was fundamentally the same, since the insulator layer for 1st and 2nd gate electrode 8a and 8d was as thin as less than 10nm at the oxide-film reduced property in the case of this operation gestalt, and it acts so that the longitudinal direction electric field may assist impregnation of a carrier. For this reason, a charge is especially injected into the carrier trap near the corner part of level difference 1a efficiently. That is, compared with the transistor structure of the 6th operation gestalt, there is an advantage which can enlarge the ratio of the charge accumulated dose to the occupancy area of a gate electrode. Moreover, charge injection efficiency becomes higher by assistance of this longitudinal direction electric field, and it becomes possible to gather that part and drawing speed.

[0106] 8th operation gestalt drawing 14 is the sectional view of the MONOS mold memory transistor concerning the 8th operation gestalt.

[0107] The point that this memory transistor differs from the MONOS mold memory transistor of the 7th operation gestalt greatly is to have not prepared the level difference in channel formation field 1a of a substrate. Therefore, the impregnation method itself is source side impregnation fundamentally. Other configurations are fundamentally [ as the 7th operation gestalt ] the same. With this operation gestalt, the gate electrode is divided and formed in the impregnation electric-field control for channel formation like the 7th operation gestalt. Moreover, in this operation gestalt, drawing speed was accelerated to 1microsec, and operating voltage is low-battery-ized to 7V. For that purpose, as compared with a well-known example, the bottom insulator layer 10 is set to 4nm or less among gate dielectric film 6, and elimination is performed by drawing out an electron in the direction of a channel.

[0108] In addition, since what is necessary is just to skip the formation process of a level difference in the manufacture approach of the 7th operation gestalt about the manufacture approach, explanation here is omitted.

[0109] Writing is made by setting up an electrical potential difference almost like the 7th operation gestalt to the memory transistor of such a configuration. For example, the program electrical potential difference which impresses the program electrical potential difference impressed to 1st gate electrode 8a at the time of writing to 5V and 2nd gate electrode 8b is set as 7V. Thus, after optimizing the bias voltage at the time of writing, it writes in by channel hot electron impregnation.

[0110] With this operation gestalt, since it divides into 1st gate electrode 8a for channel formation, and 2nd gate electrode 8b for impregnation electric-field control and the gate electrode is prepared like the 7th operation gestalt, a program electrical potential difference can be set up separately. Therefore, there is an advantage which a bias setup at the time of writing tends to optimize. Moreover, in spite of having prepared separately the gate electrode the object for channel formation, and for impregnation electric-field control, the occupancy area of the case of the 5th operation gestalt of drawing 7 and a

transistor hardly changes. Therefore, there is an advantage that the detailed memory cell suitable for high integration is realizable.

[0111] Below, other operation gestalten about the structure of a memory cell cel array configuration, a memory cell, and a memory transistor are explained.

[0112] The memory cell and memory cell array concerning the 9th operation gestalt book operation gestalt are the separation source line NOR mold with which the bit line and the source line were hierarchized. The circuitry of this NOR mold memory cell array is shown in drawing 15. Moreover, the bird's-eye view which looked at the top view of this NOR mold memory cell array to drawing 16 at drawing 17 from the cross-section side met on the B-B' line of drawing 16 is shown.

[0113] With this nonvolatile memory equipment, a bit line is hierarchized by the main bit line and the subbit line, and the source line is hierarchized by the main source line and the subsource line. The subbit line SBL1 is connected to the main bit line MBL1 through the selection transistor S11, and the subbit line SBL2 is connected to the main bit line MBL2 through the selection transistor S21. Moreover, the subsource line SSL 1 is connected to the main source line MSL1 through the selection transistor S12, and the subsource line SSL 2 is connected to the main source line MSL2 through the selection transistor S22.

[0114] Between the subbit line SBL1 and the subsource line SSL 1, parallel connection of the memory transistor M11-M1n (for example, n= 128) is carried out, and parallel connection of the memory transistor M21-M2n is carried out between the subbit line SBL2 and the subsource line SSL 2. The unit block which constitutes a memory cell array is constituted by this n memory transistor each other connected to juxtaposition, and two selection transistors (S11, S12, or S21 and S22).

[0115] Each gate of the memory transistors M11 and M21 and -- which adjoins in the direction of WORD is connected to the word line WL1. Similarly, each gate of the memory transistors M12 and M22 and -- is connected to a word line WL2, and each gate of memory transistor M1n, M2n, and -- is connected to the word line WL<sub>n</sub>. The selection transistor S11 and -- which adjoin in the direction of WORD are controlled by the selection line SG11, and the selection transistor S21 and -- are controlled by the selection line SG21. Similarly the selection transistor S12 and -- which adjoin in the direction of WORD are controlled by the selection line SG12, and the selection transistor S22 and -- are controlled by the selection line SG22.

[0116] In this detailed NOR mold cel array, as shown in drawing 17, W is formed in the front face of the semi-conductor substrate SUB p wells. Insulating separation of the W is carried out in the direction of WORD p wells by the isolation insulating layer ISO which came to embed an insulating material at a trench and has been arranged in the shape of an parallel stripe.

[0117] the isolation insulating layer ISO dissociated -- a part serves as an active region of a memory transistor p well each. On crosswise both sides in an active region, n mold impurity is introduced into high concentration in the shape of [ which kept a mutual distance ] an parallel stripe, and, thereby, the subbit lines SBL1 and SBL2 (the following, SBL, and notation) and the subsource lines SSL1 and SSL2 (the following, SSL, and notation) are formed. "The 1st impurity range" and the subsource line SSL correspond [ the subbit line SBL ] to "the 2nd impurity range." It intersects perpendicularly through an insulator layer on the subbit line SBL and the subsource line SSL, and each word lines WL1, WL2, WL3, and WL4 and -- (the following, WL, and notation) are wired at equal intervals. These word lines WL touch the interior on W and the isolation insulating layer ISO p wells through the insulator layer including a charge storage means. The subbit line part between the subbit line SBL and the subsource line SSL which touches the channel formation field by the amount of [ the part of W and / with each word line WL ] intersection becoming the channel formation field of a memory transistor functions [ the amount of drain and subsource line part ] as the source p wells.

[0118] The top face and side attachment wall of a word line WL are covered with the offset insulating layer and the sidewall insulating layer (good also in the layer insulation layer usual in this example). The bit contact BC which reaches the subbit line SBL at intervals of predetermined, and the source contact

SC which reaches the subsource line SSL are formed in these insulating layers. These contacts BC and SC are formed for every 128 memory transistors of the direction of a bit. Moreover, the main bit lines MBL1 and MBL2 which contact on the bit contact BC, --, the main source lines MSL1 and MSL2 and -- which contact on the source contact SC are formed in the shape of an parallel stripe by turns in the insulating-layer top.

[0119] The 1st highway (bit line) and the 2nd highway (source line) are hierarchized, and this detailed NOR mold cel array does not need to form the bit contact BC and the source contact SC for every memory cell. Therefore, there is no variation in the contact resistance itself fundamentally. Although they are formed for every 128 memory cells, when not performing this plug formation in self align, an offset insulating layer and the sidewall insulating layer of the bit contact BC and the source contact SC are unnecessary. That is, after depositing the usual interlayer insulation film thickly and embedding a memory transistor, opening of the contact is carried out to the usual photolithography by etching.

[0120] Since there was almost no space useless as false contact loess structure which constituted the subbit line and the subsource line from an impurity range, when each class is formed with the minimum line width F of a wafer process limitation, it is 8F2. It can manufacture in a near, very small cel area. Since an electron can be independently poured into two in one memory cell with this operation gestalt at high speed, the cel area per bit is 4F2. It becomes. Moreover, since the source line is separated, page writing is also attained. Furthermore, in order that the bit line and the source line may be hierarchized and the selection transistor S11 or S21 may separate the juxtaposition memory transistor group in a non-choosing unit block from the main bit lines MBL1 or MBL2, the capacity of the main bit line is reduced remarkably and is advantageous to improvement in the speed and low-power-izing. Moreover, by work of the selection transistors S12 or S22, a subsource line can be separated from the main source line, and can be low-capacity-ized. In addition, for the further improvement in the speed, it is good to form the subbit line SBL and the subsource line SSL in the impurity range on which silicide was stuck, and to consider the main bit line MBL and the main source line MSL as metal wiring.

[0121] The memory cell and memory cell array concerning the 10th operation gestalt of the 10th operation gestalt are the detailed NOR mold which used the self-align technique and the meandering source line. Drawing 18 is the outline top view of the NOR mold cel array concerning the 10th operation gestalt.

[0122] At this NOR mold cel array, the isolation insulating layer ISO which becomes the front face of p wells from vertical band-like a trench or LOCOS etc. is arranged in the direction of a bit (lengthwise direction of drawing 18 ) at equal intervals. the isolation insulating layer ISO -- almost -- intersecting perpendicularly -- each -- word line WLm-2, WLm-1, WLm, and WLm+1 It wires at equal intervals. The laminated structure containing this word line consists of the bottom insulator layer, a nitride, a top insulator layer, and a cascade screen of a gate electrode like the above-mentioned operation gestalt.

[0123] the active region in spacing of each isolation insulating layer -- setting -- alienation of each word line -- for example, n mold impurity is introduced into high concentration, and the source impurity range S and the drain impurity range D are formed in the tooth space by turns. That magnitude is specified in the direction of WORD (longitudinal direction of drawing 18 ) only at intervals of the isolation insulating layer ISO, and this source impurity range S and the drain impurity range D are prescribed in the direction of a bit by only WORD line spacing. Therefore, the source impurity range S and the drain impurity range D are extremely formed in homogeneity from the error of mask alignment hardly being introduced about dispersion in the magnitude and arrangement.

[0124] The perimeter of each word line only forms a sidewall insulating layer, and it is formed to the source impurity range S and the drain impurity range D, diverting the self aryne contact technique in which the contact hole for bit line connection and the contact hole for source line connection are 2 times to coincidence. And the above-mentioned process becomes unnecessary [ a photo mask ]. Therefore, as stated previously, the magnitude of the contact hole the bit line with which the magnitude of the source impurity range S and the drain impurity range D and arrangement carry out self align two-

dimensional, and are formed in a uniform top to this, or for source line connection also serves as homogeneity extremely. Moreover, the above-mentioned contact hole has the almost maximum magnitude to the area of the source impurity range S and the drain impurity range D.

[0125] Source line SL<sub>n</sub>-1 currently wired in the direction of a bit on it, SL<sub>n</sub>, and SL<sub>n</sub>+1 (the following, SL, and notation) are meanderingly arranged on the isolation insulating layer ISO and the source impurity range S, avoiding the drain impurity range D, and they are connected to each lower layer source impurity range S through the contact hole for the above-mentioned source line connection. On the source line SL, bit line BL<sub>n</sub>-1, BL<sub>n</sub>, and BL<sub>n</sub>+1 (the following, BL, and notation) are wired at equal intervals through the 2nd interlayer insulation film. This bit line BL is located above an active region, and is connected to each lower layer drain impurity range D through the contact hole for bit line connection.

[0126] By such cel pattern of a configuration, as described above, formation of the source impurity range S and the drain impurity range D cannot be easily influenced of mask alignment. Moreover, the contact hole for bit line connection and the contact hole for source line connection Since the abrupt diversion of the self aryne technique of 2 times is carried out and it is formed, a contact hole does not serve as a limit element of cel reductions of area. 6F<sub>2</sub> since source wiring etc. is made in the minimum line width F of a wafer process limitation and there is moreover almost no useless space A near, very small cel area is realizable. Being able to pour in an electron independently of [ in one memory cell ] two places, the cel area per bit is 3F<sub>2</sub> in this case. It becomes.

[0127] The 11th operation gestalt of the 11th operation gestalt is related with the non-volatile semiconductor memory (henceforth Si nano crystal mold) using Si nano crystal insulated mutually [ a large number which are embedded into gate dielectric film as a charge storage means of a memory transistor, for example, have the particle size of 10 nanometers or less ].

[0128] Drawing 19 is the sectional view showing the component structure of this Si nano crystal mold memory transistor. Si nano crystal mold nonvolatile memory of this operation gestalt differing from the previous 5th operation gestalt is that the gate dielectric film 30 of that a charge storage means is the nano crystal 32 and this operation gestalt consists of a bottom insulator layer 10 by which the Si nano crystal 32 was formed in the top face, and an oxide film 34 on it. Other configurations are the same as that of the 5th operation gestalt.

[0129] The size (diameter) is desirable, the Si nano crystal 32 is about 10nm or less, for example, 4.0nm, and each Si nano crystals are spatially separated by the oxide film 34 at intervals of about 4nm. The bottom insulator layer 10 in this example is relation close to a substrate side, and its charge storage means (Si nano crystal 32) is a little thick, and can choose it from the 5th operation gestalt suitably by within the limits from 2.6nm to 5.0nm according to a use application. Here, it considered as about 4.0nm thickness.

[0130] In manufacture of such a memory transistor of a configuration, much Si nano crystals 32 are formed on the bottom insulator layer 10 by after membrane formation of the bottom insulator layer 10 (for example, a plasma-CVD method). Moreover, an oxide film 34 is formed by LP-CVD like 7nm so that the Si nano crystal 32 may be embedded. In this LP-CVD, the mixed gas of DCS and N<sub>2</sub> O and substrate temperature may be [ material gas ] 700 degrees C. At this time, the Si nano crystal 32 is embedded at an oxide film 34. When flattening is required, it is good to newly perform planarizing processes (for example, CMP etc.). Then, the gate electrode 8 is formed and the Si nano crystal mold memory transistor concerned is completed through the process which carries out pattern NINGU of the gate cascade screen collectively.

[0131] Thus, the formed Si nano crystal 32 functions as a carrier trap discretized in the direction of a flat surface. The trap level can be presumed with a band discontinuity value with surrounding silicon oxide, and is set to about 3.1eV with the estimate. Each Si nano crystal 32 of this magnitude can hold some impregnation electrons. In addition, the Si nano crystal 32 may be made still smaller, and a single electron may be made to hold to this.

[0132] About Si nano crystal mold nonvolatile memory of such a configuration, the back tunneling model

of land KISUTO examined the data-hold property. In order to raise a data-hold property, it becomes important to make trap level deep and to enlarge distance of a charge center of gravity and the semiconductor substrate 1. Then, the simulation which used the land KISUTO model for the physical model examined the data-hold in the case of 3.1 eV trap level. Consequently, by using the deep carrier trap of 3.1 eV trap level showed that good data-hold was shown, even when the distance from a charge maintenance medium to channel formation field 1a was as near as 4.0 nm in comparison.

[0133] The 12th operation gestalt of the 12th operation gestalt is related with the non-volatile semiconductor memory (henceforth a detailed division FG mold) using the detailed assembled-die floating gate of a large number which it was embedded into the insulator layer as a charge storage means of a memory transistor, and were separated mutually.

[0134] Drawing 20 is the sectional view showing the component structure of this detailed division FG mold memory transistor. The detailed division FG mold nonvolatile memory of this operation gestalt differing from the previous 5th operation gestalt is that the gate dielectric film 40 of that a charge storage means consists of the detailed assembled-die floating gate, that the memory transistor is formed in a SOI substrate, and this operation gestalt consists of a bottom insulator layer 10 by which the detailed assembled-die floating gate 42 was formed in the top face, and an oxide film 44 on it. Other configurations are the same as that of the 5th operation gestalt. This detailed division floating gate 42 corresponds to the example of "the diameter conductor of a granule" as used in the field of [ crystal / 32 / of the previous 11th operation gestalt / Si nano ] this invention.

[0135] The SIMOX (Separation by Implanted Oxygen) substrate which carried out the ion implantation of the oxygen ion to the silicon substrate at high concentration, and formed the embedded oxide film in the part deeper than a substrate front face as a SOI substrate, the lamination substrate which formed the oxide film in one silicon substrate surface, and was made to rival other substrates are used. The SOI substrate which it was formed by such approach and shown in drawing 20 consists of a support substrate 46, a separation oxide film 48, and a silicon layer 50, and channel formation field 1a, the source impurity range 2, and the drain impurity range 4 are formed in the silicon layer 50. In addition, a glass substrate besides a semi-conductor substrate, a plastic plate, silicon on sapphire, etc. may be used for the support substrate 46.

[0136] As for the detailed division floating gate 42, the height processes the floating gate of the usual FG mold into the Pori Si dot to 8nm with a detailed diameter by about 5.0nm. although the bottom insulator layer 10 in this example is a little thicker than the 1st operation gestalt, it is markedly boiled compared with the usual FG mold, is formed thinly, and can be suitably chosen by within the limits from 2.5nm to 4.0nm according to a use application. Here, it considered as the 2.5nm thinnest thickness.

[0137] In manufacture of such a memory transistor of a configuration, after forming the bottom insulator layer 10 on a SOI substrate, the polish recon film (the last thickness: 5nm) is formed on the bottom insulator layer 10 with LP-CVD method. In this LP-CVD, the mixed gas of DCS and ammonia and substrate temperature may be [ material gas ] 650 degrees C. Next, for example, an electron-beam-exposure method is used, and a diameter processes the polish recon film into the Pori Si dot detailed [ to 8nm ]. This Pori Si dot is the detailed assembled-die floating gate 42 (charge storage means). Then, an oxide film 44 is formed by LP-CVD like 9nm in the form which embeds the detailed assembled-die floating gate 42. In this LP-CVD, the mixed gas of DCS and N<sub>2</sub>O and substrate temperature may be [ material gas ] 700 degrees C. At this time, the detailed assembled-die floating gate 42 is embedded at an oxide film 44. When flattening is required, it is good to newly perform planarizing processes (for example, CMP etc.). Then, the gate electrode 8 is formed and the detailed division FG mold memory transistor concerned is completed through the process which carries out pattern NINGU of the gate cascade screen collectively.

[0138] Thus, as a result of making a component as an experiment and evaluating a property about the floating gate being divided minutely using a SOI substrate, it checked that a good property as expected was acquired.

[0139] In the 1st described beyond the modification – the 12th operation gestalt, various deformation is possible.

[0140] Although not illustrated especially, this invention is applicable to various NOR mold cells, such as a detailed NOR mold cell which consists of cell arrays of a DINOR mold and the separation source mold shared between two source fields which are called the so-called HiCR mold and adjoin a source line.

[0141] gate dielectric film since “the charge storage means discretized superficially” in this invention contains the carrier trap formed a carrier trap and an oxide film, and near a nitride interface nitride bulk -- NO (Nitride–Oxide) the film -- this invention is applicable even if it is an MNOS mold.

[0142] This invention is applicable also to the nonvolatile memory of the embedded mold integrated on the same substrate as a logical circuit besides the nonvolatile memory of a stand-alone type. In addition, using a SOI substrate like the 12th operation gestalt overlaps the memory transistor structure of the 1st – the 11th operation gestalt, and it is applicable.

[0143]

[Effect of the Invention] According to the non-volatile semiconductor memory concerning this invention, and its drive approach, while an acceleration charge maintains momentum (a direction and magnitude), as a result of being poured into a high speed at a charge storage means, high drawing speed can be obtained efficiently. Moreover, a gate voltage setup for obtaining high drawing speed from having the 1st and 2nd gate electrode, for example as the object for channel formation and an object for impregnation electric-field control is easy. Furthermore, the application of the writing which is two or more bits with dependability and endurance high since it eliminates by electronic drawing is easy, and also has the advantage of being easy to reduce bit cost.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3. In the drawings, any words are not translated.

---

## DESCRIPTION OF DRAWINGS

---

### [Brief Description of the Drawings]

[Drawing 1] It is the circuit diagram showing the outline configuration of the source separation NOR mold memory cell array of the non-volatile semiconductor memory concerning the 1st operation gestalt.

[Drawing 2] It is the outline top view of the NOR mold memory cell array concerning the 1st operation gestalt.

[Drawing 3] It is the perspective view seen from the cross-section side which met the A-A' line by the memory cell array of drawing 2 concerning the 1st operation gestalt.

[Drawing 4] It is the sectional view showing the component structure of the MONOS mold memory transistor concerning the 1st operation gestalt.

[Drawing 5] It is the circuit diagram showing the outline configuration of the imaginary earth NOR mold memory cell array of the non-volatile semiconductor memory concerning the 2nd operation gestalt.

[Drawing 6] It is the sectional view showing the example of component structure of the memory

transistor concerning the 4th operation gestalt.

[Drawing 7] It is the sectional view showing the component structure of the MONOS memory transistor concerning the 5th operation gestalt.

[Drawing 8] It is the sectional view showing the component structure of the MONOS memory transistor concerning the 6th operation gestalt.

[Drawing 9] In manufacture of the MONOS memory transistor concerning the 6th operation gestalt, it is a sectional view after LDD field formation.

[Drawing 10] It is the sectional view after separating the gate electrode following drawing 9 and forming an insulating layer in an end face.

[Drawing 11] It is the sectional view showing the component structure of the MONOS memory transistor concerning the 7th operation gestalt.

[Drawing 12] In manufacture of the MONOS memory transistor concerning the 7th operation gestalt, it is a sectional view after membrane formation of gate dielectric film.

[Drawing 13] It is the sectional view after removing sidewall form conductive layer of one of the two following drawing 12.

[Drawing 14] It is the sectional view showing the component structure of the MONOS memory transistor concerning the 8th operation gestalt.

[Drawing 15] It is the circuit diagram showing the configuration of the NOR mold memory cell array concerning the 9th operation gestalt.

[Drawing 16] It is the top view of the NOR mold memory cell array concerning the 9th operation gestalt.

[Drawing 17] In the NOR mold memory cell array concerning the 9th operation gestalt, it is the bird's-eye view seen from the cross-section side which met the B-B' line of drawing 16.

[Drawing 18] It is the top view showing the outline configuration of the detailed NOR mold cel array concerning the 10th operation gestalt.

[Drawing 19] It is the sectional view showing the component structure of Si nano crystal mold memory transistor concerning the 11th operation gestalt.

[Drawing 20] It is the sectional view showing the component structure of the detailed division FG mold memory transistor concerning the 12th operation gestalt.

[Description of Notations]

1 -- A semi-conductor substrate or a well, 1a -- A channel formation field, 2, S -- Source impurity range (the 2nd impurity range), 4 D -- A drain impurity range (the 1st impurity range), 6, 30, 40 -- Gate dielectric film, 8 [ -- Top insulator layer, ] -- A gate electrode, 10 -- A bottom insulator layer, 12 -- A nitride, 14 32 -- 34 Si nano crystal, 44 -- An oxide film, 42 -- Detailed assembled-die floating gate, 46 -- A semi-conductor substrate, 48 -- A separation oxide film, 50 -- A silicon layer, 60--n well, 61 -- p -- a well, an ISO-- isolation insulating layer, and PW--p -- a well, and M11 -- an M22 -- memory transistor -- S11, ST0 grade [ -- A subbit line, SL1 grade / -- A source line, MSL / -- A main source line, SSL1 grade / -- A subsource line, WL1 grade / -- A word line BC / -- Bit contact SC / -- Source contact. ] -- A selection transistor, BL1 grade -- A bit line, MBL1 grade -- A main bit line, SBL

---

[Translation done.]

(51) Int.Cl.<sup>7</sup>  
 H 01 L 21/8247  
 29/788  
 29/792  
 27/115

識別記号

F I  
 H 01 L 29/78 3 7 1 5 F 0 0 1  
 27/10 4 3 4 5 F 0 8 3  
 5 F 1 0 1

テ-マコ-ト<sup>\*</sup>(参考)

審査請求 未請求 請求項の数40 OL (全 24 頁)

(21)出願番号 特願平11-361877  
 (22)出願日 平成11年12月20日 (1999.12.20)  
 (31)優先権主張番号 特願平11-277642  
 (32)優先日 平成11年9月29日 (1999.9.29)  
 (33)優先権主張国 日本 (JP)

(71)出願人 000002185  
 ソニー株式会社  
 東京都品川区北品川6丁目7番35号  
 (72)発明者 藤原 一郎  
 東京都品川区北品川6丁目7番35号 ソニ  
 一株式会社内  
 (74)代理人 100094053  
 弁理士 佐藤 隆久

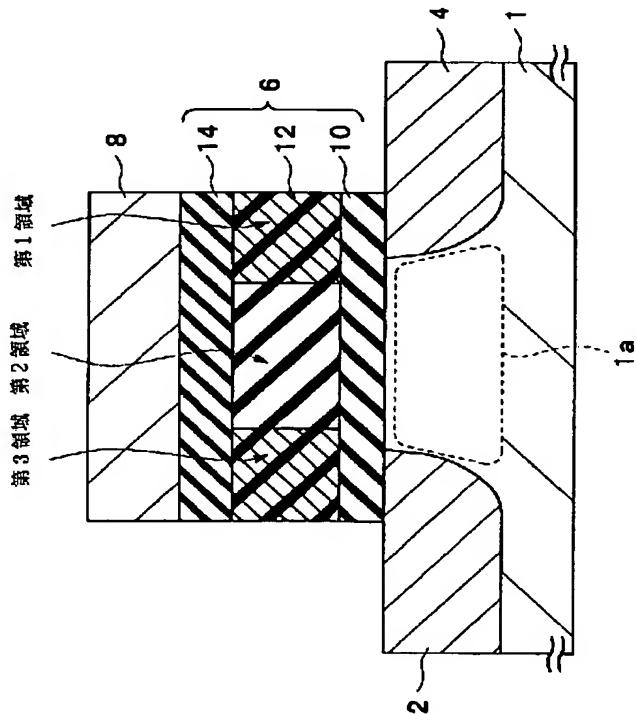
最終頁に続く

## (54)【発明の名称】 不揮発性半導体記憶装置およびその駆動方法

## (57)【要約】

【課題】 MONOS型メモリトランジスタにおいて、動作電圧を低電圧化したまま、書き込み速度を向上させる。

【解決手段】 基板1と、基板表面に設けられた半導体のチャネル形成領域1aと、チャネル形成領域1aを挟んで基板表面に形成された第1、第2不純物領域2、4と、チャネル形成領域1aに対向する面内および膜厚方向に離散化された電荷蓄積手段(キャリアトラップ)を内部に含むゲート絶縁膜6と、ゲート絶縁膜6上に設けられたゲート電極8とを備える。基板ホットエレクトロン、2次衝突電離ホットエレクトロンなど基板と垂直方向に電荷を加速させる、あるいは、チャネル形成領域1aの表面に段差1bを形成する。その結果、電荷蓄積手段が電荷加速方向に位置することとなり、注入効率が高まる。



(2)

## 【特許請求の範囲】

【請求項 1】 基板と、

当該基板の表面に設けられた半導体のチャネル形成領域と、

当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、

上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、

上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化され、かつ、動作時に電荷が加速される方向に設けられている電荷蓄積手段とを有する不揮発性半導体記憶装置。

【請求項 2】 上記電荷蓄積手段に注入される電荷がホットエレクトロンである請求項 1 に記載の不揮発性半導体記憶装置。

【請求項 3】 上記電荷の加速が、主に、上記第1不純物領域と上記ゲート電極との間の垂直電界成分により行われる請求項 1 に記載の不揮発性半導体記憶装置。

【請求項 4】 上記電荷蓄積手段は、動作時に上記電荷が注入される上記第1不純物領域側の第1領域と、電荷が注入されない第2領域とを有する請求項 3 に記載の不揮発性半導体記憶装置。

【請求項 5】 上記電荷蓄積手段は、上記第2領域を挟んで上記第1領域とチャネル方向に対峙した上記第2不純物領域側の位置に、動作時に電荷が注入される第3領域を有する請求項 4 に記載の不揮発性半導体記憶装置。

【請求項 6】 上記電荷蓄積手段に注入される電荷は、2次衝突電離により発生したホットエレクトロンである請求項 3 に記載の不揮発性半導体記憶装置。

【請求項 7】 上記電荷蓄積手段に注入される電荷は、上記チャネル形成領域の空乏層内で加速されて発生した基板ホットエレクトロンである請求項 3 に記載の不揮発性半導体記憶装置。

【請求項 8】 上記チャネル形成領域の表面に、上記第1不純物領域を上記第2不純物領域に対して相対的に低くする段差が設けられ、

上記電荷の加速が、主に、上記第1および第2不純物領域間の水平電界成分により上記段差の上部側でチャネル形成領域に沿って行われる請求項 1 に記載の不揮発性半導体記憶装置。

【請求項 9】 上記段差と上記第1不純物領域との間のゲート絶縁膜部分を中心に上記電荷蓄積手段が形成されている請求項 8 に記載の不揮発性半導体記憶装置。

【請求項 10】 上記電荷が加速される上記段差上部のチャネル方向の長さが、チャネル内電荷の平均自由行程以下である請求項 8 に記載の不揮発性半導体記憶装置。

【請求項 11】 上記電荷が加速される上記段差上部のチャネル方向の長さが、チャネル内電荷が不純物散乱の影響を受けずに走行できる所定距離以下の範囲内に設定さ

2

れている請求項 8 に記載の不揮発性半導体記憶装置。

【請求項 12】 上記電荷が加速される上記段差上部のチャネル方向の長さが、50 nm 以下である請求項 8 に記載の不揮発性半導体記憶装置。

【請求項 13】 前記電荷蓄積手段は、すくなくとも外部との間で電荷の移動がない場合に、前記チャネル形成領域に対向する面全体としての導電性を持たない請求項 1 に記載の不揮発性半導体記憶装置。

【請求項 14】 前記ゲート絶縁膜は、前記チャネル形成領域上のボトム絶縁膜と、

当該ボトム絶縁膜上の窒化膜または酸化窒化膜とを含む請求項 13 に記載の不揮発性半導体記憶装置。

【請求項 15】 前記ゲート絶縁膜は、前記チャネル形成領域上のボトム絶縁膜と、

前記電荷蓄積手段としてボトム絶縁膜上に形成され互いに絶縁された小粒径導電体とを含む請求項 13 に記載の不揮発性半導体記憶装置。

【請求項 16】 前記小粒径導電体の粒径が10ナノメータ以下である請求項 15 に記載の不揮発性半導体記憶装置。

【請求項 17】 基板と、

当該基板の表面に設けられた半導体のチャネル形成領域と、

当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、

上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、

上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有し、

上記チャネル形成領域の表面に、上記第1不純物領域を上記第2不純物領域に対し相対的に低くする段差が設けられている不揮発性半導体記憶装置。

【請求項 18】 基板と、

当該基板の表面に設けられた半導体のチャネル形成領域と、

当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、

上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、

上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有し、

上記ゲート電極は、上記チャネル形成領域の上記第1不純物領域側に、内部に電荷蓄積手段を有しない第1ゲート絶縁膜を介して積層された第1ゲート電極と、

上記チャネル形成領域の上記第2不純物領域側に、内部に電荷蓄積手段を有した第2ゲート絶縁膜を介して積層

50

(3)

3

された第2ゲート電極とを含む不揮発性半導体記憶装置。

【請求項19】上記第2ゲート電極が、上記第1ゲート電極の側壁に絶縁膜を介して形成されている請求項18に記載の不揮発性半導体記憶装置。

【請求項20】上記チャネル形成領域の表面に、上記第1不純物領域に対し上記第2不純物領域を相対的に低くする段差を備える請求項18に記載の不揮発性半導体記憶装置。

【請求項21】基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であつて、

動作時に、上記チャネル形成領域または周辺の空乏層内で電荷を加速させ、運動量を保持した状態で上記離散化された電荷蓄積手段内に注入する不揮発性半導体記憶装置の駆動方法。

【請求項22】上記加速された電荷を、上記電荷蓄積手段の上記第1不純物領域側の部分に注入する請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項23】上記第1、第2不純物領域に対するバイアス印加条件を逆にして、上記電荷蓄積手段の上記第2不純物領域側の部分に注入する請求項22に記載の不揮発性半導体記憶装置の駆動方法。

【請求項24】上記電荷の加速開始から注入までの走行距離が50nm以下である請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項25】上記電荷を、チャネル内でバリスチックに電気伝導させる請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項26】上記電荷の加速と注入を別々のゲート電極により制御する請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項27】上記ゲート電極に印加する電圧の最大値を5V以下とする請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項28】上記チャネル形成領域、上記第1、第2不純物領域、上記ゲート電極に印加する電圧の最大値を5V以下とする請求項27に記載の不揮発性半導体記憶装置の駆動方法。

【請求項29】読み出し時に、上記第1不純物領域がソースとなるように上記第1および第2不純物領域間に所

(3)

4

定の読み出しドレイン電圧を印加し、上記ゲート電極に所定の読み出しゲート電圧を印加する請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項30】上記電荷蓄積手段に部分的に注入された電荷に応じて記憶された複数ビットのデータを、上記第1、第2不純物領域間で電圧印加方向を変えて読み出す請求項23に記載の不揮発性半導体記憶装置の駆動方法。

【請求項31】消去時に、上記第1不純物領域側から注入され上記電荷蓄積手段に保持されている電荷を、直接トンネリングまたはFNトンネリングにより第1不純物領域側に引く抜く請求項22に記載の不揮発性半導体記憶装置の駆動方法。

【請求項32】消去時に、上記第1または第2不純物領域側から注入され上記電荷蓄積手段にチャネル方向の両側に分離されて保持されている電荷を、直接トンネリングまたはFNトンネリングにより個別にあるいは一括して基板側に引く抜く請求項23に記載の不揮発性半導体記憶装置の駆動方法。

【請求項33】前記電荷蓄積手段は、すくなくとも外部との間で電荷の移動がない場合に、前記チャネル形成領域に対向する面全体としての導電性を持たない請求項21に記載の不揮発性半導体記憶装置の駆動方法。

【請求項34】前記ゲート絶縁膜は、前記チャネル形成領域上のボトム絶縁膜と、当該ボトム絶縁膜上の窒化膜または酸化窒化膜とを含む請求項33に記載の不揮発性半導体記憶装置の駆動方法。

【請求項35】前記ゲート絶縁膜は、前記チャネル形成領域上のボトム絶縁膜と、前記電荷蓄積手段としてボトム絶縁膜上に形成され互いに絶縁された小粒径導電体とを含む請求項33に記載の不揮発性半導体記憶装置の駆動方法。

【請求項36】前記小粒径導電体の粒径が10ナノメータ以下である請求項35に記載の不揮発性半導体記憶装置の駆動方法。

【請求項37】基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であつて、

動作時に、上記チャネル形成領域にできたチャネル内で

(4)

5

電荷を加速させ、バリスチックな電気伝導現象を利用して上記離散化された電荷蓄積手段内に注入する不揮発性半導体記憶装置の駆動方法。

【請求項 38】基板と、

当該基板の表面に設けられた半導体のチャネル形成領域と、

当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、

上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、

上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、

動作時に、2次衝突電離により発生したホットエレクトロンを上記離散化された電荷蓄積手段内に注入する不揮発性半導体記憶装置の駆動方法。

【請求項 39】基板と、

当該基板の表面に設けられた半導体のチャネル形成領域と、

当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、

上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、

上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、

動作時に、上記チャネル形成領域にできた空乏層内で電荷を加速させ、基板ホットエレクトロンにして上記離散化された電荷蓄積手段内に注入する不揮発性半導体記憶装置の駆動方法。

【請求項 40】基板と、

当該基板の表面に設けられた半導体のチャネル形成領域と、

当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、

上記チャネル形成領域上に形成されたゲート絶縁膜と、当該ゲート絶縁膜上に形成されたゲート電極と、

上記チャネル形成領域に対向した面内および膜厚方向に離散化されて上記ゲート絶縁膜内に形成され、動作時にチャネルホットエレクトロン、バリスチックホットエレクトロン、2次衝突電離ホットエレクトロン、基板ホットエレクトロンまたはバンド間トンネル電流に起因したホットエレクトロンが注入される電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、

消去時に、上記第1および/または第2不純物領域側か

6

ら注入され上記電荷蓄積手段にチャネル方向の一方側または両側に保持されている電荷を、直接トンネリングまたはF Nトンネリングにより個別にあるいは一括して基板側に引く抜く不揮発性半導体記憶装置の駆動方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、メモリトランジスタのチャネル形成領域とゲート電極との間のゲート絶縁膜の内部に、平面的に離散化された電荷蓄積手段（例えば、MONOS型やMOS型における窒化膜内の電荷トラップ、トップ絶縁膜と窒化膜との界面近傍の電荷トラップ、或いは小粒径導電体等）を有し、当該電荷蓄積手段に対し電荷（電子またはホール）を電気的に注入して蓄積し又は引き抜くことを基本動作とする不揮発性半導体記憶装置と、その駆動方法とに関する。

【0002】

【従来の技術】不揮発性半導体メモリは、大容量で小型の情報記録媒体として期待されているが、近年、情報ネットワークの広帯域化とともにネットワークの伝送速度（たとえば搬送波周波数：100MHz）と同等の書き込み速度が要求されるようになってきている。このため、不揮発性半導体メモリに対し、スケーリング性が良好で従来の  $100\mu\text{s e c}$  /セルの書き込み速度より1桁またはそれ以上の書き込み速度の向上が要求されている。

【0003】不揮発性半導体メモリは、電荷を保持する電荷蓄積手段が平面的に連続したFG (Floating Gate)型のほかに、電荷蓄積手段が平面的に離散化された、例えばMONOS (Metal-Oxide-Nitride-Oxide Semiconductor)型などがある。

【0004】MONOS型不揮発性半導体メモリでは、電荷保持を主体的に担っている窒化膜 [Si<sub>x</sub>N<sub>y</sub> (0 < x < 1, 0 < y < 1)] 膜中またはトップ絶縁膜と窒化膜との界面のキャリアトラップが空間的に（即ち、面方向および膜厚方向に）離散化して拡がっているために、電荷保持特性が、トンネル絶縁膜厚のほかに、Si<sub>x</sub>N<sub>y</sub>膜中のキャリアトラップに捕獲される電荷のエネルギー的および空間的な分布に依存する。

【0005】このトンネル絶縁膜に局的にリーク電流40 パスが発生した場合、FG型では多くの電荷がリークパスを通ってリークして電荷保持特性が低下しやすいのに対し、MONOS型では、電荷蓄積手段が空間的に離散化されているため、リークパス周辺の局的な電荷がリークパスを通して局的にリークするに過ぎず、記憶素子全体の電荷保持特性が低下しにくい。このため、MONOS型においては、トンネル絶縁膜の薄膜化による電荷保持特性の低下の問題はFG型ほど深刻ではない。したがって、ゲート長が極めて短い微細メモリトランジスタにおけるトンネル絶縁膜のスケーリング性は、MONOS型の方がFG型よりも優れている。また、平面的に

(5)

7

離散化したキャリアトラップの分布平面に対し電荷が局所的に注入された場合、その電荷はFG型のように平面内および膜厚方向に拡散することなく保持される。

【0006】MONOS型不揮発性メモリで微細メモリセルを実現するにはディスターブ特性の改善が重要であり、そのためにはトンネル絶縁膜を通常の膜厚(1.6nm~2.0nm)より厚く設定する必要が生じている。

【0007】

【発明が解決しようとする課題】ところが、従来のMONOS型不揮発性メモリにおいて、トンネル絶縁膜を厚膜化したり動作電圧を低減することは書き込み速度向上にとっては益々不利な状況となっている。このため、従来のMONOS型等の不揮発性メモリでは、信頼性(たとえば、データ保持特性、リードディスターブ特性あるいはデータ書換え特性など)を十分に満足させた場合、書き込み速度は100μsecが限界である。

【0008】書き込み速度だけを考えると、チャネルホットエレクトロン(CHE)を用いた注入方式のほうが、チャネル全面FNトンネリングより高速化しやすい。ところが、ドレイン端でCHEを発生させる通常のCHE注入方式では、注入効率が $1 \times 10^{-6}$ と十分ではない。また、注入効率を上げるために、CHEをソース側から注入するソースサイド注入型MONOSトランジスタが報告されたが(IEEE Electron Device Letter 19, 1998, pp153)、このソースサイド注入型MONOSトランジスタでは、動作電圧が書き込み時12V、消去時14Vと高いうえ、リードディスターブ特性およびデータ書換え特性などの信頼性が十分でない。

【0009】このように、従来のMONOS型等の不揮発性メモリでは、書き込み速度の向上と、動作電圧の低電圧化および信頼性確保がトレードオフの関係にある。このトレードオフの克服は、高速な大容量不揮発性メモリを開発するうえで、さらに近年盛んに開発が進められているシステムLSIにおいてロジック回路との混載の観点から、重要な課題となっていた。

【0010】本発明の目的は、MONOS型など平面的に離散化されたキャリアトラップ等に電荷を蓄積させて基本動作するメモリトランジスタにおいて、リードディスターブ特性などの信頼性を良好に維持し、また動作電圧を低電圧化したまま、書き込み速度を向上することが可能な構造を有する不揮発性半導体記憶装置と、その駆動方法を提供することである。

【0011】

【課題を解決するための手段】本発明の第1の観点に係る不揮発性半導体記憶装置は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、

10

20

30

40

40

50

8

当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化され、かつ、動作時に電荷が加速される方向に設けられている電荷蓄積手段とを有する。この電荷が主に加速される方向は、基板に水平方向と垂直方向の何れでもよい。

【0012】この第1の観点に係る不揮発性半導体記憶装置は、離散化された電荷蓄積手段を電荷が加速される方向に設けていることから、注入の際に電荷の運動量が維持されやすい。このため、電荷蓄積手段に対し効率よく電荷が注入される。

【0013】本発明の第2の観点に係る不揮発性半導体記憶装置は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有し、上記チャネル形成領域の表面に、上記第1不純物領域を上記第2不純物領域に対し相対的に低くする段差が設けられている。

【0014】好適に、上記段差と上記第1不純物領域との間のゲート絶縁膜部分を中心に上記電荷蓄積手段が形成されている。好適に、上記電荷が加速される上記段差上部のチャネル方向の長さ(たとえば、チャネル形成領域の第2不純物領域端から段差までの長さ)が、チャネル内電子の平均自由行程以下である。あるいは、この長さは、チャネル内電子が不純物散乱の影響を受けずに走行できる所定距離以下の範囲内、たとえば50nm以下に設定されている。

【0015】一般に、ホットエレクトロン注入を用いて書き込みを行う不揮発性半導体記憶装置では、ソースとドレイン間に所定のバイアス電圧が印加され、またゲート電極に所定の書き込み電圧が印加されたときに、形成されたチャネル内にソースから供給された電荷(この場合、電子)が電界加速される。その加速によりドレイン近傍で高エネルギーを得た電荷(ホットエレクトロン)は、ゲート電極による電界に引き付けられて電荷蓄積手段に注入される。

【0016】本発明の第2の観点に係る不揮発性半導体記憶装置は、基板表面に段差を設け、段差の低い側でONO膜などの内部に電荷蓄積手段(キャリアトラップ)を有するゲート絶縁膜を設けている。したがって、ドレイン近傍で発生した高エネルギー電荷(たとえば、ホットエレクトロン)は、その運動量(方向と大きさ)を維持しながら殆ど運動エネルギーを失うことなく効率良く、しかも高速にキャリアトラップに注入される。第2の観点に係る不揮発性半導体記憶装置は、第1の観点に係る不揮発性半導体記憶装置の水平電荷加速の一形態を

(6)

9

示す。この電荷が加速される段差上部のチャネル方向の長さたとえば50nm以下とすると、電荷がチャネル内を不純物散乱等の影響を殆ど受けずにパリスチックに伝導する。したがって、電荷注入の効率、速度が更に高まる。

【0017】本発明の第3の観点に係る不揮発性半導体記憶装置は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向する面内および膜厚方向に離散化されている電荷蓄積手段とを有し、上記ゲート電極は、上記チャネル形成領域の上記第1不純物領域側に、内部に電荷蓄積手段を有しない第1ゲート絶縁膜を介して積層された第1ゲート電極と、上記チャネル形成領域の上記第2不純物領域側に、内部に電荷蓄積手段を有した第2ゲート絶縁膜を介して積層された第2ゲート電極とを含む。好適に、第2ゲート電極が第1ゲート電極の側壁に絶縁膜を介して形成されている。

【0018】この第3の観点に係る不揮発性半導体記憶装置では、第1不純物領域側の第1ゲート電極、第2不純物領域側の第2ゲート電極が別々に設けられている。したがって、加速時のゲート電圧印加条件と、注入時のゲート電圧印加条件を別々に設定でき、それだけ、注入効率を極大化して書き込みを高速化するための制御がしやすい。第1ゲート電極側のゲート絶縁膜（第1ゲート絶縁膜）内に電荷蓄積手段を形成しないでよいので、第1ゲート絶縁膜を薄くして加速電界を大きくできる。また、加速電界を一定とすれば、印加電圧を下げることができる。一方、第2ゲート電極をサイドウォール形とした場合、セル面積が小さくて済む。

【0019】本発明の第4の観点に係る不揮発性半導体記憶装置の駆動方法は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向する面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、動作時に、上記チャネル形成領域または周辺の空乏層内で電荷を加速させ、運動量を保持した状態で上記離散化された電荷蓄積手段内に注入する。好適に、書き込みの際に、電荷蓄積手段の第1不純物領域側と第2不純物領域側に独立に部分的な電荷注入を行う。具体的には、電圧印加条件を逆にして書き込みを行う。読み出しでは、好適に、読み出すべき情報に対応した電荷がソース側となるよう

10

に、第1、第2不純物領域間の電圧印加方向を決め、所定の読み出しドレイン電圧を印加し、また、ゲート電極に所定の読み出しゲート電圧を印加する。複数ビットの読み出しでは、このような読み出しを、第1、第2不純物領域間の電圧印加方向を逆にして行う。消去では、上記第1および/または第2不純物領域側から注入され上記電荷蓄積手段にチャネル方向の一方側または両側に保持されている電荷を、直接トンネリングまたはFNTNトンネリングにより個別にあるいは一括して基板側に引く抜く。

【0020】本発明の第5の観点に係る不揮発性半導体記憶装置の駆動方法は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、動作時に、上記チャネル形成領域にできたチャネル内で電荷を加速させ、パリスチックな電気伝導現象を利用して上記離散化された電荷蓄積手段内に注入する。

【0021】本発明の第6の観点に係る不揮発性半導体記憶装置の駆動方法は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、動作時に、2次衝突電離により発生したホットエレクトロンを上記離散化された電荷蓄積手段内に注入する。

【0022】本発明の第7の観点に係る不揮発性半導体記憶装置の駆動方法は、基板と、当該基板の表面に設けられた半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に設けられたゲート絶縁膜と、当該ゲート絶縁膜上に設けられたゲート電極と、上記ゲート絶縁膜内で、上記チャネル形成領域に対向した面内および膜厚方向に離散化されている電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、動作時に、上記チャネル形成領域にできた空乏層内で電荷を加速させ、基板ホットエレクトロンにして上記離散化された電荷蓄積手段内に注入する。

【0023】本発明の第8の観点に係る不揮発性半導体記憶装置の駆動方法は、基板と、当該基板の表面に設け

50

(7)

11

られ半導体のチャネル形成領域と、当該チャネル形成領域を挟んで基板表面に形成され、動作時にソースまたはドレインとなる第1および第2不純物領域と、上記チャネル形成領域上に形成されたゲート絶縁膜と、当該ゲート絶縁膜上に形成されたゲート電極と、上記チャネル形成領域に対向した面内および膜厚方向に離散化されて上記ゲート絶縁膜内に形成され、動作時にチャネルホットエレクトロン、パリスチックホットエレクトロン、2次衝突電離ホットエレクトロン、基板ホットエレクトロンまたはバンド間トンネル電流に起因したホットエレクトロンが注入される電荷蓄積手段とを有する不揮発性半導体記憶装置の駆動方法であって、消去時に、上記第1および／または第2不純物領域側から注入され上記電荷蓄積手段にチャネル方向の一方側または両側に保持されている電荷を、直接トンネリングまたはF Nトンネリングにより個別にあるいは一括して基板側に引く抜く。

【0024】第4～第7の観点に係る不揮発性半導体記憶装置の駆動方法では、たとえば、段差を介したチャネルホットエレクトロン注入、パリスチック伝導を用いた電荷注入、2次衝突電離ホットエレクトロン注入または基板ホットエレクトロン注入を用いるために、電荷注入の効率がよく、使用電圧が低くても十分に高速な書き込みが可能である。本発明では、注入方式を適宜選択することで、電荷が加速される方向を基板に水平方向のほかに垂直方向にすることもでき、また、平面的に離散化された電荷蓄積手段に対し電荷を部分的に注入することもできる。また、第8の観点に係る不揮発性半導体記憶装置の駆動方法では、電荷を基板側にトンネリングに引き抜くことにより消去を行うことから、従来のように消去動作中にボトム絶縁膜内をホールが多量に移動することができない。なお、本発明は、ゲート絶縁膜内でボトム絶縁膜上に窒化膜または酸化窒化膜を含むMONOS型またはMNOS型等、あるいはゲート絶縁膜内でボトム絶縁膜上に互いに絶縁された小粒径導電体を含む小粒径導電体型に好適である。

【0025】

#### 【発明の実施の形態】第1実施形態

図1は、本発明の実施形態に係るソース線分離NOR型の不揮発性半導体メモリのメモリセルアレイの概略構成を示す回路図である。

【0026】この不揮発性メモリ装置では、NOR型メモリセルアレイの各メモリセルがメモリトランジスタ1個で構成されている。図1に示すように、メモリトランジスタM11～M22が行列状に配置され、これらトランジスタ間がワード線、ピット線および分離型ソース線によって配線されている。すなわち、ピット方向に隣接するメモリトランジスタM11およびM12の各ドレインがピット線BL1に接続され、各ソースがソース線SL1に接続されている。同様に、ピット方向に隣接するメモリトランジスタM21およびM22の各ドレインが

50

12

ピット線BL2に接続され、各ソースがソース線SL2に接続されている。また、ワード方向に隣接するメモリトランジスタM11とM21の各ゲートがワード線WL1に接続され、同様に、ワード方向に隣接するメモリトランジスタM12とM22の各ゲートがワード線WL2に接続されている。メモリセルアレイ全体では、このようなセル配置およびセル間接続が繰り返されている。

【0027】図2は、第1実施形態に係る微細NOR型セルアレイの概略平面図である。また、図3は、図2のA-A'線に沿った断面側から見た鳥瞰図である。

【0028】この微細NOR型メモリセルアレイでは、図3に示すように、n型またはp型の半導体基板SUB(nウエルまたはpウエルでも可)の表面にトレンチまたはLOCOSなどから素子分離絶縁層ISOが形成されている。素子分離絶縁層ISOは、図2に示すように、ピット方向(図2の縦方向)に長い平行ストライプ状に配置されている。素子分離絶縁層ISOにほぼ直交して、各ワード線WL1, WL2, WL3, WL4, …が等間隔に配線されている。このワード線は、後述するように、ボトム絶縁膜、窒化膜、トップ絶縁膜からなるゲート絶縁膜と、ゲート電極とを積層させて構成されている。

【0029】各素子分離絶縁層ISOの間隔内の能動領域において、各ワード線の離間スペースに、基板101と逆導電型の不純物が高濃度に導入されてソース不純物領域(第2不純物領域)Sとドレイン不純物領域(第1不純物領域)Dとが交互に形成されている。このソース不純物領域Sとドレイン不純物領域Dは、その大きさがワード方向(図2の横方向)には素子分離絶縁層ISOの間隔のみで規定され、ピット方向にはワード線間隔のみで規定される。したがって、ソース不純物領域Sとドレイン不純物領域Dは、その大きさと配置のばらつきに関しマスク合わせの誤差が殆ど導入されないことから、極めて均一に形成されている。

【0030】ワード線の上部および側壁は、絶縁層で覆われている。すなわち、ワード線WL1, WL2, …の上部に同じパターンにてオフセット絶縁層が配置され、オフセット絶縁層、その下のゲート電極(ワード線)およびゲート絶縁膜からなる積層パターンの両側壁に、サイドウォール絶縁層が形成されている。このオフセット絶縁層およびサイドウォール絶縁層により、各ワード線同士のスペース部分に、ワード線に沿って細長い自己整合コンタクトホールが開口されている。

【0031】ソース不純物領域Sまたはドレイン不純物領域Dに一部重なるように、自己整合コンタクトホール内に導電性材料が互い違いに埋め込まれ、これによりピットコンタクトBCおよびソースコンタクトSCが形成されている。これらコンタクトBC, SCの形成では、自己整合コンタクトホール内を埋め込むように導電材料を堆積し、その上に、エッチングマスク用のレジストパ

(8)

13

ターンを形成する。このとき、レジストパターンを自己整合コンタクトホールの幅より一回り大きくし、また、一部を素子分離絶縁層 I SO に重ねる。そして、このレジストパターンをマスクとしてレジストパターン周囲の導電材料をエッチングにより除去する。これにより、2種類のコンタクト BC, SC が同時に形成される。

【0032】図示しない絶縁膜でコンタクト周囲の凹部が埋め込まれている。この絶縁膜上を、ピットコンタクト BC 上に接触するピット線 BL 1, BL 2, …と、ソースコンタクト SC 上に接触するソース線 SL 1, SL 2, …が交互に、平行ストライプ状に形成されている。

【0033】この微細NOR型セルアレイは、そのピット線またはソース線に対するコンタクト形成が、自己整合コンタクトホールの形成と、プラグの形成により達成される。自己整合コンタクトホールの形成では、ワード線との絶縁分離が達成されるとともに、ソース不純物領域 S またはドレイン不純物領域 D の表面が均一に形成される。そして、ピットコンタクト BC およびソースコンタクト SC の形成は、この自己整合コンタクトホール内のソース不純物領域 S またはドレイン不純物領域 D の表面に対して行う。したがって、各プラグの基板接触面は、そのピット方向のサイズがほぼ自己整合コンタクトホールの形成により決められ、その分、コンタクト面積のバラツキは小さい。

【0034】ピットコンタクト BC またはソースコンタクト SC と、ワード線との絶縁分離が容易である。すなわち、ワード線形成時に一括してオフセット絶縁層を形成しておき、その後、絶縁膜の成膜と、全面エッチング（エッチバック）を行うだけでサイドウォール絶縁層が形成される。また、ピットコンタクト BC とソースコンタクト SC、さらに、ピット線とソース線が同一階層の導電層をパターンニングして形成されるため、配線構造が極めて簡素であり、工程数も少なく、製造コストを低く抑えるのに有利な構造となっている。しかも、無駄な空間が殆どないことから、各層の形成をウエハプロセス限界の最小線幅 F で行った場合、 $8 F^2$  に近い非常に小さいセル面積で製造できる。

【0035】図4は、本実施形態に係るMONOS型メモリトランジスタの素子構造を示す断面図である。

【0036】図4中、符号1はn型またはp型の導電型を有するシリコンウェハ等の半導体基板SUBまたはウエル（以下、基板という）、1aはチャネル形成領域、2および4は該メモリトランジスタのソース不純物領域 S およびドレイン不純物領域 D を示す。本発明で“チャネル形成領域”とは、表面側内部に電子または正孔が導電するチャネルが形成される領域をいう。本例の“チャネル形成領域”は、基板内でソース不純物領域 2 およびドレイン不純物領域 4 に挟まれた部分が該当する。

【0037】ソース不純物領域 2 およびドレイン不純物領域 4 は、チャネル形成領域 1a と逆導電型の不純物を

(8)

14

高濃度に基板 1 に導入することにより形成された導電率が高い領域であり、種々の形態がある。通常、ソース不純物領域 2 及びドレイン不純物領域 4 のチャネル形成領域 1a に臨む基板表面位置に、LDD (Lightly Doped Drain) と称する低濃度領域を具備させことが多い。

【0038】チャネル形成領域 1a 上に、ゲート絶縁膜 6 を介してメモリトランジスタのゲート電極 8 が積層されている。ゲート電極 8 は、一般に、p型またはn型の不純物が高濃度で導入されて導電化されたポリシリコン (doped poly-Si)、又はdoped poly-Si と高融点金属シリサイドとの積層膜からなる。このゲート電極 8 のチャネル方向の長さ（ゲート長）は、0.25 μm 以下、たとえば0.18 μm 程度である。

【0039】本実施形態におけるゲート絶縁膜 6 は、下層から順に、ボトム絶縁膜 10、窒化膜 12、トップ絶縁膜 14 から構成されている。ボトム絶縁膜 10 は、たとえば、酸化膜を形成し、これを窒化処理して用いる。ボトム絶縁膜 10 の膜厚は、使用用途に応じて 2.0 nm から 5.0 nm の範囲内で決めることができ、ここでは 5.0 nm に設定されている。

【0040】窒化膜 12 は、例えば 5.0 nm の窒化シリコン (Si<sub>x</sub>N<sub>y</sub> (0 < x < 1, 0 < y < 1)) 膜から構成されている。この窒化膜 12 は、たとえば減圧 CVD (LPCVD) により作製され、膜中にキャリアトラップが多く含まれ、ブルーフレンケル型 (PF型) の電気伝導特性を示す。

【0041】トップ絶縁膜 14 は、窒化膜 12 との界面近傍に深いキャリアトラップを高密度に形成する必要があり、このため、例えば成膜後の窒化膜を熱酸化して形成される。また、トップ絶縁膜 14 を HTO (High Temperature chemical vapor deposited Oxide) 法により形成した SiO<sub>2</sub> 膜としてもよい。トップ絶縁膜 14 が CVD で形成された場合は熱処理によりこのトラップが形成される。トップ絶縁膜 14 の膜厚は、ゲート電極 8 からのホールの注入を有効に阻止してデータ書換可能な回数の低下防止を図るために、最低でも 3.0 nm、好ましくは 3.5 nm 以上が必要である。ここでは、トップ絶縁膜厚を 3.5 nm とする。

【0042】このような構成のメモリトランジスタの製造においては、まず、用意した半導体基板に素子分離絶縁層 I SO、ウエル W を形成し、しきい値電圧調整用のイオン注入等を必要に応じて行った後に、ゲート絶縁膜 6 を成膜する。具体的に、たとえば、短時間高温熱処理法 (RTO法) により 1000°C, 10 sec の熱処理を行い、酸化シリコン膜（ボトム絶縁膜 10）を形成する。つぎに、ボトム絶縁膜 10 上に LPCVD 法により窒化シリコン膜（窒化膜 12）を、最終膜厚が 5 nm となるように、これより厚めに堆積する。この CVD は、たとえば、ジクロロシラン (DCS) とアンモニアを混合したガスを用い、基板温度 650°C で行う。形成

(9)

15

した窒化シリコン膜表面を熱酸化法により酸化して、たとえば3.5 nmの酸化シリコン膜（トップ絶縁膜14）を形成する。この熱酸化は、たとえばH<sub>2</sub>O雰囲気中において炉温度950°Cで行う。これにより、トラップレベル（窒化シリコン膜の伝導帯からのエネルギー差）が2.0 eV以下の程度の深いキャリアトラップが約1～2×10<sup>13</sup>/cm<sup>2</sup>の密度で形成される。また、窒化膜12が1 nmに対し熱酸化シリコン膜（トップ絶縁膜14）が約1.6 nm形成され、この割合で下地の窒化膜厚が減少し、窒化膜12の最終膜厚が5 nmとなる。

【0043】ゲート電極8となる導電膜、オフセット絶縁層（不図示）の積層膜を積層させる。そして、このゲート絶縁膜6、導電膜およびオフセット絶縁層の積層膜を一括して同一パターンにて加工する。形成した積層パターンと自己整合的にソースおよびドレイン不純物領域2, 4を、イオン注入法により形成する。

【0044】続いて、図3のメモリセルアレイ構造とするために、サイドウォール絶縁層とともに自己整合コンタクトホールを形成し、自己整合コンタクトホールにより表出するソースおよびドレイン不純物領域S, D（ソースおよびドレイン不純物領域2, 4）上にピットコンタクトBCおよびソースコンタクトSCを形成する。その後、これらコンタクト周囲を層間絶縁膜で埋め込み、層間絶縁膜上にピット線BL1, …およびソース線SL1, …を形成した後、必要に応じて行う層間絶縁層を介した上層配線の形成およびオーバーコート成膜とパッド開口工程等を経て、当該不揮発性メモリセルアレイを完成させる。

【0045】つぎに、このような構成の不揮発性メモリの書き込み時のバイアス設定例および動作について、メモリトランジスタM11にデータを書き込む場合を例に説明する。本実施形態では、書き込みを2次衝突電離により発生したホットエレクトロンを用いて行う。2次衝突電離ホットエレクトロン注入では、ドレイン近傍で発生した基板電流となる正孔がドレイン近傍の空乏層を横切って基板に注入されるときに、空乏層内で電界からエネルギーを受け取って電子、正孔対を発生させ、このうち発生した電子が電界によって主に垂直方向に加速され、電荷蓄積手段に注入される。

【0046】書き込み時に、図4において、たとえば、ソース不純物領域2に0V、ドレイン不純物領域4に3.3V、ゲート電極8に5V、ウエル1に3Vを印加する。この条件下、ドレイン不純物領域4から注入され近傍の空乏層に入った正孔がシリコン原子に衝突し、この衝突によって当該空乏層内で電子、正孔対が発生する。そのうち正孔はより電位が低い基板側へ消散するが、電子は、ゲート電極と基板間の電界で加速されて空乏層を上方に向かって加速される。この加速によって電子はホットエレクトロンとなり、その一部がボトム絶縁

16

膜10のエネルギー障壁を乗り越えて窒化膜12中の電荷蓄積手段（キャリアトラップ）に注入される。このとき、衝突を引き起こす正孔が注入される側、すなわち図4における第1領域に注入電子の分布が局在する。このため、当該メモリトランジスタM11のしきい値電圧が消去状態の場合に書き込み状態まで上昇し、書き込みがなされる。選択ワード線に連なる複数のメモリセルに対し、この書き込みと書き込み禁止を電圧印加条件を変えて適宜設定することで、ページ書き込みが可能である。

【0047】読み出しでは、バイアス値を書き込み状態に応じてチャネルが形成される程度に変更する。たとえば、ドレイン不純物領域4を接地した状態で、ソース不純物領域2に1.5V、ゲート電極8に2Vを印加する。これにより、ページ読み出しの場合、電荷蓄積手段の第1領域に電子が注入されていない消去状態のメモリトランジスタではチャネルが形成され、電荷蓄積手段の第1領域に電子が注入された書き込み状態のメモリトランジスタではチャネルが形成されない。したがって、ピット線BL1, …には、そのチャネル形成の有無に応じた電位変化が現出する。この電位変化をセンスアンプで検出すると、ページ内の記憶データが一括して読み出される。

【0048】消去では、チャネル全面から、あるいはドレイン不純物領域4側からFNトンネリングまたは直接トンネリングを用いて電荷を引く抜くことにより行う。

【0049】これに対し、1セル内に2ビットのデータを記憶するために、電荷蓄積手段のソース不純物領域側にも同様な書き込みを行う場合、この2回目の書き込みでは、ソースとドレインの印加電圧を1回目と逆とする。これにより、1回目と同様な2次衝突電離がソース不純物領域側で起こり、発生した電子が電荷蓄積手段のソース不純物領域側の領域（第3領域）に注入される。2ビットとも書き込み状態のセルでは、電荷蓄積手段の第1領域にホットエレクトロンが注入されて保持され、これと独立に、第3領域にホットエレクトロンが注入され保持されている。つまり、電荷蓄積手段の第1領域と第3領域との間にはホットエレクトロンが注入されない第2領域が介在するので、この2ビット情報に対応したホットエレクトロンは確実に峻別される。

【0050】第3領域に保持された電荷の読み出しは、第1領域の電荷の場合とはソースとドレイン間の電圧方向を逆にして行う。これにより、2ビットのデータを独立に読み出すことができる。消去も、前記した第1領域側の消去と、ソース不純物領域2とドレイン不純物領域4の印加電圧を逆にして行う。なお、チャネル全面で消去を行う場合は、ソースおよびドレイン不純物領域2, 4を共に基板と同電位とすることで、第1領域側と第3領域側のデータが一括消去される。

【0051】書き込み状態、消去状態のメモリトランジスタの電流-電圧特性について検討した。この結果、ド

(10)

17

レイン電圧1.5Vでの非選択セルからのオフリード電流は約1nAであった。この場合の読み出し電流は10μA以上であるため、非選択セルの誤読み出しが生じることはない。したがって、ゲート長0.18μmのMONOS型メモリトランジスタにおいて読み出し時のパンチスルーナイフのマージンは十分あることが分かった。また、ゲート電圧1.5Vでのリードディスターブ特性も評価したが、3×10<sup>8</sup>sec以上の時間経過後でも読み出しが可能であることが分かった。

【0052】データ書換え回数は、キャリアトラップが空間的に離散化されているために良好で、1×10<sup>6</sup>回を満足することが分かった。また、データ保持特性は1×10<sup>6</sup>回のデータ書換え後で85℃、10年を満足した。

【0053】以上より、ゲート長0.18μmのMONOS型不揮発性メモリトランジスタとして十分な特性が得られていることを確かめることができた。

【0054】本実施形態における2次衝突電離ホットエレクトロン注入法では、離散化された電荷蓄積手段に対し、比較的低いドレイン電流でホットエレクトロンの注入が可能である。したがって、書き込み時のパンチスルーナイフが通常のホットエレクトロン注入と比較して抑制され、その結果、ゲート長のスケーリングが容易であるという利点がある。また、電荷が主に基板と垂直方向に加速されるため、加速電荷の運動量が維持されたまま注入が行われるため、通常のCHE注入方式に比べ、電荷の注入効率が高い。

#### 【0055】第2実施形態

第2実施形態は、仮想接地NOR型の不揮発性メモリ装置に関する。仮想接地型は、大別すると、スプリットゲート型とAMG(Alternate Metal Virtual Ground)型がある。スプリットゲート型は選択トランジスタの機能を実質的にメモリトランジスタ内に具備させることで書き込みディスターブを防止する。AMG型は半導体不純物領域からなる拡散層配線を1本おきにメタル配線に接続させ、たとえば、メタル配線をビット線として、メタル配線間の拡散層配線をソース線として選択することで隣接セルの書き込みディスターブを防止する。

【0056】図5は、仮想接地NOR型のうちAMG型のメモリセルアレイ構成を示す回路図である。このメモリセルアレイでは、ソース線が隣接メモリセル間で共通化されている。この共通化されたソース線は隣りのメモリセルを動作させるときはビット線として機能する。したがって、このメモリセルアレイでは、ビット方向の配線は全て“ビット線”と称する。各ビット線BL1～BL3は、半導体の不純物領域からなる拡散層配線である。その1本おき、たとえばビット線BL1とBL3が、図示しないビットコンタクトを介して上層のメタル配線に接続されている。

【0057】このようにセルアレイ構造が異なる以外、

18

MONOS型メモリトランジスタ構造、および、書き込み、読み出しあり消去動作の基本は、第1実施形態と同様である。仮想接地NOR型の場合、隣接するセルのビット線でソース線を代替えて用いることから、分離ソース線型と比べビット方向のサイズが縮小でき、セル面積6F<sup>2</sup>が達成できる。1つのメモリセルに物理的に2ビットのデータを記録することができ、この場合、ビット当たりのセル面積は実効的に3F<sup>2</sup>となる。

#### 【0058】第3実施形態

10 本実施形態では、チャネルホットエレクトロン(CHE)注入書き込みと、チャネル全面トンネル消去を行う。書き込みは通常のCHE注入であり、ここでの詳細は省略する。

【0059】メモリトランジスタを図4と同じ構成とし、電荷蓄積手段の第1領域に保持された電子をチャネル全面から直接トンネリングを用いて引き抜く場合、ゲート電極8に0V、ドレイン不純物領域4に8V、ソース不純物領域2をオープン、ウェル1に8Vの電圧を印加する。これにより、電荷蓄積手段の第1領域に保持されていた電子が、基板側に引き抜かされることで、セル消去が行われる。このとき、消去速度は1msec程度であった。また、第3領域の電荷を消去するときは、第1領域側の消去と、ソース不純物領域2とドレイン不純物領域4の印加電圧を逆にして行う。さらに、第1、第3領域に電荷が保持されている場合、チャネル全面で消去を行う際に、ソースおよびドレイン不純物領域2、4を共に基板と同電位とすることで、第1領域側と第3領域側のデータを一括消去してもよい。

【0060】この消去方法では、ホットホール注入などに比べ動作中にボトム絶縁膜10を通過するホール量が格段に少なく、そのためボトム絶縁膜10が劣化し難く、信頼性および耐久性(たとえば、エンディランス特性)に優れる。

#### 【0061】第4実施形態

第4実施形態では、基板ホットエレクトロン注入現象を利用した高速書き込み方法について、MONOS型メモリトランジスタを例として説明する。

【0062】基板ホットエレクトロン注入方式では、ソースとドレインを同電位で印加し、基板バイアス電圧を印加して、ゲート電極下に厚い空乏層を形成した状態で、空乏化しているチャネル形成領域に電荷(電子)を注入する。注入された電子は、空乏層内で加速されて絶縁膜の障壁エネルギー以上のエネルギーを得て、絶縁膜を越えて平面的に離散化された電荷蓄積手段に注入され、これにより書き込みがなされる。

【0063】本実施形態に係るMONOS型メモリトランジスタの第1の構成は、第1実施形態における図4と同じである。本実施形態では、基本構成は第1実施形態と同じであるが、ここでのゲート絶縁膜6の各構成膜150, 12, 14は、たとえば3/5/3.5nmとす

(11)

19

る。また、このゲート絶縁膜6の作製およびメモリセルの作製は、前記した実施形態と同様な装置、プロセス条件を用いて行う。

【0064】つぎに、メモリセル動作について説明する。ソース不純物領域2とウエル1間のpn接合に、当該pn接合が順方向と逆方向とで交互にバイアスされるACバイアス電圧を印加する。この場合、pn接合が順バイアスのときはpn接合より電子がpウエル1に注入される。ウエル内のチャネル形成領域は、基板バイアスにより空乏化させてある。このため、pn接合より注入された電子は、ゲートに印加された正バイアス電圧により加速されボトム絶縁膜10の障壁ポテンシャルを越えて電荷蓄積手段である窒化膜12中のキャリアトラップにホットキャリア注入され、これにより書き込みがなされる。

【0065】たとえば、ドレインをオープン、ウエル電位を0Vで保持した状態で、ゲート電極8に正バイアス電圧5V、ACバイアス電圧としてソース不純物領域2に対しパルス電圧を正方向に0.7V、逆方向にV<sub>CC</sub>で印加した。その場合、ホットキャリア注入が高速に行われ、動作電圧5V以下で書き込み時間1μsec以下が実現できた。また、信頼性については、従来のFNトンネル注入方式のMONOS型メモリセルと同等の特性が得られた。

【0066】電荷蓄積手段が平面的に離散化されているため、基板ホットエレクトロンの注入条件を変えることにより、窒化膜12全面への注入のほかに、窒化膜12への部分的な注入が可能である。たとえば、上記のようにソース側からの注入では電荷蓄積手段のソース側部分（第1領域）への電荷注入がなされるが、逆に、ソースオープンとしてドレイン側から電荷を注入することにより電荷蓄積手段のドレイン側部分（第3領域）に電荷注入が可能である。図4の場合、第1領域と第2領域との間に、電荷が注入されない第2領域が形成されるため、2ビットの情報の峻別が可能である。また、消去の際の電荷の引き抜きも局所的に行われ、電荷注入されていない第2領域のしきい値電圧は変化しないため、過剰消去が防止され、消去状態のしきい値電圧の収束性が高いという利点がある。その一方、ソースとドレインに対し、大きな電圧値でACバイアス電圧を同じ位相で印加することにより、図4の第1、第3領域がつながり、電荷蓄積手段全体に電荷を注入することができる。

【0067】図6に、本実施形態に係るMONOS型メモリトランジスタの第2の構造例を示す。図6に示すメモリトランジスタは、2重ウエル構造となっている。すなわち、p型半導体基板1にnウエル60が形成され、nウエル60内にpウエル61が形成され、このpウエル61内にメモリトランジスタが形成されている。他の基本構成は、図4と同じである。ただし、本例のゲート絶縁膜6の各構成膜10、12、14の膜厚仕様は、

20

3.5/5/4nmとした。

【0068】nウエル60とpウエル61で構成されるpn接合を順バイアスすることにより、空乏化されたチャネル形成領域に電子を注入し、この電子をゲート電極側に電界加速した後、基板ホットエレクトロン注入で窒化膜12内のキャリアトラップに注入する。

【0069】たとえば、ドレインオープンとした状態で、ゲート電圧5V、ソース電圧V<sub>CC</sub>を設定し、nウエル60とpウエル61間に順方向に0.7Vを印加した。これにより、動作電圧5V以下で書き込み時間1μsec以下が実現できた。また、メモリセルの信頼性については、従来のFNトンネル注入方式のMONOS型メモリセルと同等の特性が得られた。

【0070】基板ホットエレクトロン注入も、2次衝突電離の場合と同様に、動作電圧が5V以下と低いうえ、加速が主として基板に垂直に行われることから注入効率が高いという利点がある。

#### 【0071】第5実施形態

第5実施形態では、チャネル形成領域に段差を設け、この段差からCHE注入を行う。図7に、第5実施形態に係るMONOSメモリトランジスタの断面図を示す。

【0072】このメモリトランジスタが、図4に示す第1実施形態に係るメモリトランジスタと異なる点は、チャネル形成領域1a部分の基板表面に段差1bが設けられていることである。段差1bは5~50nm程度、たとえば10nmの高さを有し、相対的にソース不純物領域2側が高く、ドレイン不純物領域4側が低くなるように形成されている。他の構成は、第1実施形態と同様である。ただし、ここではゲート絶縁膜6の各構成膜10、12、14の膜厚仕様を、3.5/8.0/3.5nmとする。

【0073】このような構成のメモリトランジスタの製造において、第1実施形態と異なる点は段差1bを形成することであるが、その詳細は後述の実施形態で述べる。

【0074】つぎに、このような構成の不揮発性メモリに対するCHE注入書き込み動作について、図1のメモリトランジスタM11にデータを書き込む場合を例に説明する。

【0075】書き込み時に、必要に応じて書き込みインヒビット電圧を設定した後、プログラム電圧を印加する。たとえば、選択ビット線BL1に5Vを印加し、非選択ワード線WL2、非選択ソース線SL2および非選択ビット線BL2に、基板電位0Vのときは所定の電圧、例えば3Vを印加する。また、選択ソース線SL1は、接地電位0Vで保持する。この状態で、選択ワード線WL1に、基板電位0Vのときは所定の電圧、例えば7Vを印加する。

【0076】この書き込み条件下、ソース不純物領域2とドレイン不純物領域4との間のチャネル形成領域表面

(12)

21

に反転層（チャネル）が形成され、チャネル内にソース不純物領域2側から電子が注入されて、電界加速される。加速された電子が水平チャネル端付近でホットエレクトロンとなり、その一部がボトム絶縁膜10のポテンシャル壁を越えて高エネルギー注入され、ゲート絶縁膜6内のキャリアトラップに捕獲される。このため、当該メモリトランジスタM11のしきい値電圧が消去状態から書き込み状態まで上昇し、書き込みがなされる。

【0077】本実施形態に係るトランジスタ構造では、チャネル形成領域1aの途中に段差1bが設けられ、段差1bの底部側にゲート絶縁膜6が延在する。したがって、ゲート絶縁膜6の膜厚仕様に合わせて段差1bの高さを最適化しておくと、電子の走行方向の延長線上またはゲート電極8の印加電圧による電界で曲げられる軌跡上に、キャリアトラップの分布中心を位置させることができる。つまり、本実施形態に係るメモリトランジスタの書き込みでは、電界加速された電子の運動量（大きさと方向）をほぼ維持した状態でのキャリアトラップへの直接注入が可能となる。したがって、注入時のエネルギー損失が従来より低減し、高速で効率がよい電荷注入が可能で、その結果として、書き込み速度の高速化が達成される。

【0078】従来のチャネルホットエレクトロン注入では、エネルギー的に励起された後の散乱によりチャネルから飛び出した電荷を電界でキャリアトラップ側に引き寄せていたため、その注入効率が $1 \times 10^{-6}$ であり100万個に1個と低い割合でしかなかった。これに対し、本実施形態に係るトランジスタ構造では、チャネルホットエレクトロン注入の場合に注入効率が1桁以上改善され、 $10 \mu\text{sec}$ 以下の書き込み速度が達成できた。

【0079】さらに、電子が加速される部分のチャネル長、すなわちソース不純物領域2の端から段差1bまでの距離をたとえば50nm以下に短くすると、この距離がチャネル電子の平均自由行程と同程度またはそれ以下になるので、電界加速された電子がバリスチックにチャネル内を電気伝導する。このバリスチック電子は、不純物散乱等の影響を受けずに弾道的に高速にキャリアトラップに注入されるため、さらに注入効率が高くでき、データ書き込みが高速化される。

【0080】書き込み速度を現状維持とする場合、または、ある程度の高速化で十分である場合には、このようなトランジスタ構造および書き込み方法の適用によって、チャネル形成領域（基板またはウエル）とゲート電極間の印加電圧を10V以下にすることができる。ゲート電極と基板またはウエルとの電圧印加を正電源と負電源で分割して行う場合、動作電圧を絶対値で5V以下に低減することができる。

【0081】電荷の注入効率が上がるため、書き込み時のチャネル電流を低減する余地が生まれる。したがって、従来のチャネルホットエレクトロン注入では高電圧

22

回路の電流駆動能力の限界により難しかった、たとえば同一ワード線に接続された多数のメモリセルに対して一括して行う並列書き込みが可能になった。

【0082】なお、消去については、通常のごとく、チャネル全面からのFNトンネリングを用いた電荷引き抜きによりブロック一括して行った場合、その消去速度が $100 \mu\text{sec}$ 程度であった。

#### 【0083】第6実施形態

図8は、第6実施形態に係るMONOS型メモリトランジスタの断面図である。

【0084】このメモリトランジスタが第5実施形態のMONOS型メモリトランジスタと大きく異なる点は、ゲート電極がソース側の第1ゲート電極8aと、ドレイン側の第2ゲート電極8bに分割されていることである。第1ゲート電極8aは、電子が加速される部分、すなわち段差1b上部に対向して設けられ、第2ゲート電極8bは、主に段差1b底部に対向して設けられている。書き込み時において、第1ゲート電極8aは主に電荷を加速させるチャネルの生成を制御し、第2ゲート電極8bは主に電荷を注入する電界の制御を行う。ゲート絶縁膜6を構成する膜のうち、窒化膜12とトップ絶縁膜14が第1ゲート電極8a側に延在していない。つまり、第1ゲート電極8aとチャネル形成領域1aとの間にはボトム絶縁膜10のみ介在している。これに対し、第2ゲート電極8bとチャネル形成領域1aとの間には、第5実施形態と同様に3層構造のゲート絶縁膜6が介在する。第1および第2ゲート電極8a, 8bの隙間は絶縁層9aで埋め込まれ、また、第1および第2ゲート電極8a, 8bの外側面には、それぞれサイドウォール絶縁層9bが形成されている。

【0085】このサイドウォール絶縁層9bの形成前後のイオン注入により、ソース・ドレイン不純物領域2, 4とLDD領域2a, 4aからなるLDD構造の不純物領域が基板1の表面に形成されている。また、ドレイン側のLDD領域4a端から段差1bにかけて、チャネル形成領域1aの表面に薄くp型の不純物領域3が形成されている。なお、このp型の不純物領域3および上記LDD領域2a, 4aは、必須の構成でない。

【0086】図9(A)～図10(E)に、このMONOS型メモリトランジスタの製造方法の一例を断面図により示す。

【0087】素子分離絶縁層およびウエル等を半導体基板の表面に形成した後、図9(A)において、基板1の表面に段差1bを形成する。この段差1bの形成では、図示のように、基板表面の一部をマスク層、たとえばレジストR1により覆った後、ドライエッチングによりレジストR1に保護されていないシリコン表面を所定深さエッチングする。続いて、同じレジストR1をマスク層としたイオン注入を行い、形成した段差1bの底部および側部にp型不純物領域3を形成する。

(13)

23

【0088】レジストR1を除去後、図9 (B) では、ボトム絶縁膜10、窒化膜12およびトップ絶縁膜14を、第1実施形態と同様な方法によって、順次成膜する。その後、段差1bから、その底部側の一部を覆うマスク層、たとえばレジストR2のパターンを形成した後、ドライエッチングによりレジストR2周囲のトップ絶縁膜14および窒化膜12を除去する。

【0089】レジストR2を除去後、図9 (C) では、ゲート電極となる導電膜、たとえばdoped poly-Si膜8cを堆積し、先の工程でパターンニングしたトップ絶縁膜14および窒化膜12上から段差上部の所定部分を覆うマスク層、たとえばレジストR3を形成する。このレジストR3をマスクとして、レジストR3周囲のdoped poly-Si膜8cを除去する。その後、同じレジストR3をマスク層としたイオン注入を行い、doped poly-Si膜8cより外側の基板1の表面に、低濃度のn型LDD領域2a、4aを形成する。このイオン注入の前または後で、レジストR3をマスク層としたエッチングにより、周辺部のボトム絶縁膜10を除去する。

【0090】レジストR3の除去後、図10 (D) では、doped poly-Si膜8cの中央部分を横切る開口部を有するレジストR4を形成する。レジストR4をマスクとして、その開口部から表出するdoped poly-Si膜8cの中央部分を除去する。これにより、第1ゲート電極8aと、第2ゲート電極8bが分離して形成される。

【0091】レジストR4の除去後、図10 (E) では、全面に、たとえば酸化シリコン系の絶縁膜を堆積し、第1および第2ゲート電極8a、8bの周囲を絶縁物で覆い、かつ、両電極間の隙間を絶縁物で埋め込む。この状態で全面異方性エッチング(エッチバック)を行うことにより、第1および第2ゲート電極8a、8b間の絶縁層9aと、サイドウォール絶縁層9bを同時に形成する。

【0092】図8に示すように、第1および第2ゲート電極8a、8b、絶縁層9a、9bを自己整合マスクとしてn型不純物を高濃度にイオン注入し、ソースおよびドレイン不純物領域2、4を形成する。その後は、第1実施形態と同様な諸工程を経て、メモリセルアレイを完成させる。

【0093】このような構成のメモリトランジスタに対して、第5実施形態とほぼ同じ様にして電圧を設定することで、書き込みがなされる。このとき、本実施形態ではチャネル形成用の第1ゲート電極8aと注入電界制御用の第2ゲート電極8bとに分けてゲート電極が設けられていることから、プログラム電圧を別々に設定することができる。したがって、書き込み時のバイアス設定が最適化しやすい利点がある。

【0094】たとえば、書き込み時に、第1ゲート電極8aに印加するプログラム電圧を3V、第2ゲート電極8bに印加するプログラム電圧を5Vに設定する。この

24

ように書き込み時のバイアス電圧を最適化したうえで、バリスチック電子伝導による直接注入を行うと、書き込み速度を1μsec以下、たとえば100nsec程度まで高速化することが可能となった。

【0095】第7実施形態

図11は、第7実施形態に係るMONOS型メモリトランジスタの断面図である。

【0096】このメモリトランジスタが第6実施形態のMONOS型メモリトランジスタと大きく異なる点は、電荷注入電界を制御するための第2ゲート電極8dをサイドウォール形としたことである。これにより、トランジスタの占有面積を第2実施形態の場合よりかなり小さくすることが可能となった。サイドウォール形の第2ゲート電極8dの形成にともない、第1および第2ゲート電極8a、8d間の絶縁層に、ゲート絶縁膜6を構成する窒化膜12とトップ絶縁膜14を用いている。その他の構成は、第6実施形態と基本的に同じである。

【0097】なお、第2ゲート電極8dをサイドウォール形としたことにより、ドレイン不純物領域4と段差1bとの距離が第2実施形態の場合より近くなり、その分、チャネルの形成が容易になった。したがって、図11では、チャネル形成領域1aの表面にp型不純物領域が形成されていないが、もちろん、第6実施形態と同様にp型不純物領域3を設けてもよい。また、第6実施形態と同様、ソースおよびドレイン不純物領域2、4の内側にそれぞれLDD領域を設けてもよい。

【0098】図12 (A) ~図13 (D) に、このMONOS型メモリトランジスタの製造方法の一例を断面図により示す。

【0099】まず、図12 (A) に示すように、基板表面に段差1bを第6実施形態と同様な方法により形成する。つぎに、ボトム絶縁膜10およびゲート電極となる導電膜を成膜し、図示しないレジストなどをマスクとしたエッチングにより、ゲート電極となる導電膜をパターンニングする。これにより、段差1bの上部の所定位置に第1ゲート電極8aが形成される。

【0100】レジストを除去後、図12 (B) では、窒化膜12およびトップ絶縁膜14を、第1実施形態と同様な方法によって、順次成膜する。

【0101】その後、図13 (C) において、ゲート電極となる導電膜を全面に厚く堆積し、これをエッチバックする。これにより、第1ゲート電極8aの両側壁に、窒化膜12およびトップ絶縁膜14を介してサイドウォール形の導電層8d、8eが形成される。

【0102】図13 (D) では、まず、この導電層8d、8eを自己整合マスクとして、その周囲に表出するトップ絶縁膜14、窒化膜12およびボトム絶縁膜10を順次除去する。段差底部側の導電層(第2ゲート電極)8dおよび第1ゲート電極8a上を覆ってマスク層、たとえばレジストR5を形成する。レジストR5を

(14)

25

マスクにエッチングを行って、片方の導電層 8 e を除去し、続いて、トップ絶縁膜 1 4、窒化膜 1 2 およびボトム絶縁膜 1 0 を順次除去する。

【0103】レジスト R 5 を除去後、図 1 1 に示すように、第 1 および第 2 ゲート電極 8 a、8 d、絶縁膜 1 2、1 4 を自己整合マスクとして n 型不純物を高濃度にイオン注入し、ソースおよびドレイン不純物領域 2、4 を形成する。その後は、第 1 実施形態と同様な諸工程を経て、メモリセルアレイを完成させる。

【0104】このような構成のメモリトランジスタに対する書き込みでは、チャネル形成用と注入電界制御用のゲート電極を別々に設けているにもかかわらず、図 7 の第 5 実施形態の場合とトランジスタの占有面積が殆ど変わらない。したがって、高集積化に適した微細メモリセルが実現できるという利点がある。

【0105】書き込み時の電圧設定方法は、第 6 実施形態と基本的に同じであるが、本実施形態の場合、第 1 および第 2 ゲート電極 8 a、8 d 間の絶縁膜が酸化膜換算値で 10 nm 未満と薄いので、第 1 ゲート電極 8 a の印加電圧に応じた横方向の電界強度が高く、その横方向電界がキャリアの注入をアシストするように作用する。このため、とくに段差 1 a のコーナー部分に近いキャリアトラップに効率よく電荷が注入される。つまり、第 6 実施形態のトランジスタ構造に比べ、ゲート電極の占有面積に対する電荷蓄積量の比率を大きくできる利点がある。また、この横方向電界のアシストによって電荷注入効率がより高くなり、その分、書き込み速度を上げることが可能となる。

#### 【0106】第 8 実施形態

図 1 4 は、第 8 実施形態に係る MONOS 型メモリトランジスタの断面図である。

【0107】このメモリトランジスタが第 7 実施形態の MONOS 型メモリトランジスタと大きく異なる点は、基板のチャネル形成領域 1 a に段差を設けていない点にある。したがって、注入方式そのものは基本的にはソースサイド注入である。その他の構成は、第 7 実施形態と基本的に同じである。本実施形態では、第 7 実施形態と同様に、ゲート電極をチャネル形成用との注入電界制御用とに分けて形成している。また、本実施形態においては、書き込み速度を 1 μsec に高速化し、かつ、動作電圧を 7 V に低電圧化している。そのためには、公知例と比較して、ゲート絶縁膜 6 のうちボトム絶縁膜 1 0 を 4 nm 以下にし、かつ、消去をチャネル方向に電子を引き抜くことにより行う。

【0108】なお、製造方法については、第 7 実施形態の製造方法において段差の形成工程を省略すればよいので、ここでの説明は省略する。

【0109】このような構成のメモリトランジスタに対して、第 7 実施形態とほぼ同じ様にして電圧を設定することで、書き込みがなされる。たとえば、書き込み時

26

に、第 1 ゲート電極 8 a に印加するプログラム電圧を 5 V、第 2 ゲート電極 8 b に印加するプログラム電圧を 7 V に設定する。このように書き込み時のバイアス電圧を最適化したうえで、チャネルホットエレクトロン注入により書き込みを行う。

【0110】本実施形態では、第 7 実施形態と同様、チャネル形成用の第 1 ゲート電極 8 a と注入電界制御用の第 2 ゲート電極 8 b とに分けてゲート電極が設けられていることから、プログラム電圧を別々に設定することができる。したがって、書き込み時のバイアス設定が最適化しやすい利点がある。また、チャネル形成用と注入電界制御用のゲート電極を別々に設けているにもかかわらず、図 7 の第 5 実施形態の場合とトランジスタの占有面積が殆ど変わらない。したがって、高集積化に適した微細メモリセルが実現できるという利点がある。

【0111】以下に、メモリセルセルアレイ構成、メモリセルおよびメモリトランジスタの構造に関する他の実施形態を説明する。

#### 【0112】第 9 実施形態

20 本実施形態に係るメモリセルおよびメモリセルアレイは、ピット線およびソース線が階層化された分離ソース線 NOR 型である。図 1 5 に、この NOR 型メモリセルアレイの回路構成を示す。また、図 1 6 に、この NOR 型メモリセルアレイの平面図を、図 1 7 に、図 1 6 の B-B' 線に沿った断面側から見た鳥瞰図を示す。

【0113】この不揮発性メモリ装置では、ピット線が主ピット線と副ピット線に階層化され、ソース線が主ソース線と副ソース線に階層化されている。主ピット線 M-BL 1 に選択トランジスタ S 1 1 を介して副ピット線 S-BL 1 が接続され、主ピット線 M-BL 2 に選択トランジスタ S 2 1 を介して副ピット線 S-BL 2 が接続されている。また、主ソース線 M-SL 1 に選択トランジスタ S 1 2 を介して副ソース線 S-SL 1 が接続され、主ソース線 M-SL 2 に選択トランジスタ S 2 2 を介して副ソース線 S-SL 2 が接続されている。

【0114】副ピット線 S-BL 1 と副ソース線 S-SL 1 との間に、メモリトランジスタ M 1 1 ~ M 1 n (たとえば、n = 1 2 8) が並列接続され、副ピット線 S-BL 2 と副ソース線 S-SL 2 との間に、メモリトランジスタ M 2 1 ~ M 2 n が並列接続されている。この互いに並列に接続された n 個のメモリトランジスタと、2 つの選択トランジスタ (S 1 1 と S 1 2、又は、S 2 1 と S 2 2) とにより、メモリセルアレイを構成する単位ブロックが構成される。

【0115】ワード方向に隣接するメモリトランジスタ M 1 1、M 2 1、…の各ゲートがワード線 WL 1 に接続されている。同様に、メモリトランジスタ M 1 2、M 2 2、…の各ゲートがワード線 WL 2 に接続され、また、メモリトランジスタ M 1 n、M 2 n、…の各ゲートがワード線 WL n に接続されている。ワード方向に隣接する

(15)

27

選択トランジスタ S 1 1, …は選択線 SG 1 1 により制御され、選択トランジスタ S 2 1, …は選択線 SG 2 1 により制御される。同様に、ワード方向に隣接する選択トランジスタ S 1 2, …は選択線 SG 1 2 により制御され、選択トランジスタ S 2 2, …は選択線 SG 2 2 により制御される。

【0116】この微細NOR型セルアレイでは、図17に示すように、半導体基板SUBの表面にpウエルWが形成されている。pウエルWは、トレンチに絶縁物を埋め込んでなり平行ストライプ状に配置された素子分離絶縁層ISOにより、ワード方向に絶縁分離されている。

【0117】素子分離絶縁層ISOにより分離された各pウエル部分が、メモリトランジスタの能動領域となる。能動領域内の幅方向両側で、互いの距離をおいた平行ストライプ状にn型不純物が高濃度に導入され、これにより、副ビット線SBL1, SBL2(以下、SBLと表記)および副ソース線SSL1, SSL2(以下、SSLと表記)が形成されている。副ビット線SBLが“第1不純物領域”、副ソース線SSLが“第2不純物領域”に該当する。副ビット線SBLおよび副ソース線SSL上に絶縁膜を介して直交して、各ワード線WL1, WL2, WL3, WL4, …(以下、WLと表記)が等間隔に配線されている。これらのワード線WLは、内部に電荷蓄積手段を含む絶縁膜を介してpウエルW上および素子分離絶縁層ISO上に接している。副ビット線SBLと副ソース線SSLとの間のpウエルWの部分と、各ワード線WLとの交差部分がメモリトランジスタのチャネル形成領域となり、そのチャネル形成領域に接する副ビット線部分がドレイン、副ソース線部分がソースとして機能する。

【0118】ワード線WLの上面および側壁は、オフセット絶縁層およびサイドウォール絶縁層(本例では、通常の層間絶縁層でも可)により覆われている。これら絶縁層には、所定間隔で副ビット線SBLに達するビットコンタクトBCと、副ソース線SSLに達するソースコンタクトSCとが形成されている。これらのコンタクトBC, SCは、たとえば、ビット方向のメモリトランジスタ128個ごとに設けられている。また、絶縁層上を、ビットコンタクトBC上に接触する主ビット線MBL1, MBL2, …と、ソースコンタクトSC上に接触する主ソース線MSL1, MSL2, …が交互に、平行ストライプ状に形成されている。

【0119】この微細NOR型セルアレイは、第1共通線(ビット線)および第2共通線(ソース線)が階層化され、メモリセルごとにビットコンタクトBCおよびソースコンタクトSCを形成する必要がない。したがって、コンタクト抵抗自体のバラツキは基本的でない。ビットコンタクトBCおよびソースコンタクトSCは、たとえば128個のメモリセルごとに設けられるが、このプラグ形成を自己整合的に行わないときは、オフセット

28

絶縁層およびサイドウォール絶縁層は必要ない。すなわち、通常の層間絶縁膜を厚く堆積してメモリトランジスタを埋め込んだ後、通常のフォトリソグラフィとエッチングによりコンタクトを開口する。

【0120】副ビット線、副ソース線を不純物領域で構成した疑似コンタクトレス構造として無駄な空間が殆どないことから、各層の形成をウエハプロセス限界の最小線幅Fで行った場合、 $8F^2$ に近い非常に小さいセル面積で製造できる。本実施形態では1つのメモリセル内の2か所に、独立に電子を高速で注入できるため、ビット当たりのセル面積は $4F^2$ となる。また、ソース線を分離しているため、ページ書き込みも可能となる。さらに、ビット線とソース線が階層化されており、選択トランジスタS11又はS21が非選択の単位ブロックにおける並列メモリトランジスタ群を主ビット線MBL1またはMBL2から切り離すため、主ビット線の容量が著しく低減され、高速化、低消費電力化に有利である。また、選択トランジスタS12またはS22の働きで、副ソース線を主ソース線から切り離して、低容量化することができる。なお、更なる高速化のためには、副ビット線SBLおよび副ソース線SSLをシリサイドを張りつけた不純物領域で形成し、主ビット線MBLおよび主ソース線MSLをメタル配線とするとよい。

#### 【0121】第10実施形態

第10実施形態に係るメモリセルおよびメモリセルアレイは、自己整合技術と蛇行ソース線を用いた微細NOR型である。図18は、第10実施形態に係るNOR型セルアレイの概略平面図である。

【0122】このNOR型セルアレイでは、pウエルの表面に縦帯状のトレンチまたはLOCOSなどからなる素子分離絶縁層ISOが等間隔でビット方向(図18の縦方向)に配置されている。素子分離絶縁層ISOにはほぼ直交して、各ワード線WL<sub>m-2</sub>, WL<sub>m-1</sub>, WL<sub>m</sub>, WL<sub>m+1</sub>が等間隔に配線されている。このワード線を含む積層構造は、前述の実施形態と同様に、ボトム絶縁膜、窒化膜、トップ絶縁膜及びゲート電極の積層膜から構成されている。

【0123】各素子分離絶縁層の間隔内の能動領域において、各ワード線の離間スペースに、例えばn型不純物が高濃度に導入されてソース不純物領域Sとドレイン不純物領域Dとが交互に形成されている。このソース不純物領域Sとドレイン不純物領域Dは、その大きさがワード方向(図18の横方向)には素子分離絶縁層ISOの間隔のみで規定され、ビット方向にはワード線間隔のみで規定される。したがって、ソース不純物領域Sとドレイン不純物領域Dは、その大きさと配置のばらつきに関しまスク合わせの誤差が殆ど導入されないことから、極めて均一に形成されている。

【0124】各ワード線の周囲は、サイドウォール絶縁層を形成するだけで、ソース不純物領域Sとドレイン不

(16)

29

純物領域Dとに対し、ピット線接続用のコンタクトホールとソース線接続用のコンタクトホールとが2度のセルフアラインコンタクト技術を同時に転用しながら形成される。しかも、上記プロセスはフォトマスクが不要となる。したがって、先に述べたようにソース不純物領域Sとドレイン不純物領域Dの大きさや配置が均一な上に、これに対して2次元的に自己整合して形成されるピット線またはソース線接続用のコンタクトホールの大きさも極めて均一となる。また、上記コンタクトホールはソース不純物領域Sとドレイン不純物領域Dの面積に対し、ほぼ最大限の大きさを有している。

【0125】その上でピット方向に配線されているソース線 $S_{Ln-1}$ ,  $S_{Ln}$ ,  $S_{Ln+1}$ （以下、 $S_L$ と表記）は、ドレイン不純物領域Dを避けながら素子分離絶縁層ISO上とソース不純物領域S上に蛇行して配置され、上記ソース線接続用のコンタクトホールを介して、下層の各ソース不純物領域Sに接続されている。ソース線 $S_L$ 上には、第2の層間絶縁膜を介してピット線 $B_{Ln-1}$ ,  $B_{Ln}$ ,  $B_{Ln+1}$ （以下、 $B_L$ と表記）が等間隔で配線されている。このピット線 $B_L$ は、能動領域上方に位置し、ピット線接続用のコンタクトホールを介して、下層の各ドレイン不純物領域Dに接続されている。

【0126】このような構成のセルパターンでは、上記したように、ソース不純物領域Sとドレイン不純物領域Dの形成がマスク合わせの影響を受けにくく、また、ピット線接続用のコンタクトホールとソース線接続用のコンタクトホールが、2度のセルフアライン技術を一括転用して形成されることから、コンタクトホールがセル面積縮小の制限要素とはならず、ウエハプロセス限界の最小線幅Fでソース配線等ができる、しかも、無駄な空間が殆どないことから、 $6F^2$ に近い非常に小さいセル面積が実現できる。1つのメモリセル内の2か所に独立に電子を注入することができ、この場合、ピット当たりのセル面積は $3F^2$ となる。

#### 【0127】第11実施形態

第11実施形態は、メモリトランジスタの電荷蓄積手段としてゲート絶縁膜中に埋め込まれれば10ナノメータ以下の粒径を有する多数の互いに絶縁されたSiナノ結晶を用いた不揮発性半導体記憶装置（以下、Siナノ結晶型という）に関する。

【0128】図19は、このSiナノ結晶型メモリトランジスタの素子構造を示す断面図である。本実施形態のSiナノ結晶型不揮発性メモリが、先の第5実施形態と異なるのは、電荷蓄積手段がナノ結晶32であること、および、本実施形態のゲート絶縁膜30が、Siナノ結晶32が上面に形成されたボトム絶縁膜10と、その上の酸化膜34とから構成されていることである。その他の構成は、第5実施形態と同様である。

【0129】Siナノ結晶32は、そのサイズ（直径）が、好ましくは10nm以下、例えば4.0nm程度で

(17)

30

あり、個々のSiナノ結晶同士が酸化膜34で空間的に、例えば4nm程度の間隔で分離されている。本例におけるボトム絶縁膜10は、電荷蓄積手段（Siナノ結晶32）が基板側に近いこととの関係で、第5実施形態よりやや厚く、使用用途に応じて2.6nmから5.0nmまでの範囲内で適宜選択できる。ここでは、4.0nm程度の膜厚とした。

【0130】このような構成のメモリトランジスタの製造では、ボトム絶縁膜10の成膜後、例えばプラズマCVD法でボトム絶縁膜10の上に、多数のSiナノ結晶32を形成する。また、Siナノ結晶32を埋め込むように、酸化膜34を、例えば7nmほどLP-CVDにより成膜する。このLP-CVDでは、原料ガスがDCSと $N_2O$ の混合ガス、基板温度が例えば700°Cとする。このときSiナノ結晶32は酸化膜34に埋め込まれる。平坦化が必要な場合は、新たに平坦化プロセス（例えばCMP等）を行うとよい。その後、ゲート電極8を成膜し、ゲート積層膜を一括してパターンニングする工程を経て、当該Siナノ結晶型メモリトランジスタを完成させる。

【0131】このように形成されたSiナノ結晶32は、平面方向に離散化されたキャリアトラップとして機能する。そのトラップレベルは、周囲の酸化シリコンとのバンド不連続値で推定可能で、その推定値では約3.1eV程度とされる。この大きさの個々のSiナノ結晶32は、数個の注入電子を保持できる。なお、Siナノ結晶32を更に小さくして、これに単一電子を保持させてもよい。

【0132】このような構成のSiナノ結晶型不揮発性メモリについて、ランドキストのバックトンネリングモデルによりデータ保持特性を検討した。データ保持特性を向上させるためには、トラップレベルを深くして、電荷重心と半導体基板1との距離を大きくすることが重要となる。そこで、ランドキストモデルを物理モデルに用いたシミュレーションにより、トラップレベル3.1eVの場合のデータ保持を検討した。この結果、トラップレベル3.1eVの深いキャリアトラップを用いることにより、電荷保持媒体からチャネル形成領域1aまでの距離が4.0nmと比較的に近い場合でも良好なデータ保持を示すことが分かった。

#### 【0133】第12実施形態

第12実施形態は、メモリトランジスタの電荷蓄積手段として絶縁膜中に埋め込まれ互いに分離した多数の微細分割型フローティングゲートを用いた不揮発性半導体記憶装置（以下、微細分割FG型という）に関する。

【0134】図20は、この微細分割FG型メモリトランジスタの素子構造を示す断面図である。本実施形態の微細分割FG型不揮発性メモリが、先の第5実施形態と異なるのは、電荷蓄積手段が微細分割型フローティングゲートからなること、メモリトランジスタがSOI基板

(17)

31

に形成されていること、および、本実施形態のゲート絶縁膜40が、微細分割型フローティングゲート42が上面に形成されたボトム絶縁膜10と、その上の酸化膜44とから構成されていることである。その他の構成は、第5実施形態と同様である。この微細分割フローティングゲート42は、先の第11実施形態のSiナノ結晶32とともに本発明でいう“小粒径導電体”的具体例に該当する。

【0135】SOI基板としては、酸素イオンをシリコン基板に高濃度にイオン注入し基板表面より深い箇所に埋込酸化膜を形成したSIMOX (Separation by Implanted Oxygen) 基板や、一方のシリコン基板表面に酸化膜を形成し他の基板と張り合わせた張り合せ基板などが用いられる。このような方法によって形成され図20に示したSOI基板は、支持基板46、分離酸化膜48およびシリコン層50とから構成され、シリコン層50内に、チャネル形成領域1a、ソース不純物領域2およびドレイン不純物領域4が設けられている。なお、支持基板46は、半導体基板のほか、ガラス基板、プラスチック基板、サファイア基板等を用いてもよい。

【0136】微細分割フローティングゲート42は、通常のFG型のフローティングゲートを、その高さが例えば5.0nm程度で、直径が例えば8nmまでの微細なポリSiドットに加工したものである。本例におけるボトム絶縁膜10は、第1実施形態よりやや厚いが、通常のFG型に比べると格段に薄く形成され、使用用途に応じて2.5nmから4.0nmまでの範囲内で適宜選択できる。ここでは、最も薄い2.5nmの膜厚とした。

【0137】このような構成のメモリトランジスタの製造では、SOI基板上にボトム絶縁膜10を成膜した後、例えばLPCVD法で、ボトム絶縁膜10の上にポリシリコン膜（最終膜厚：5nm）を成膜する。このLPCVDでは、原料ガスがDCSとアンモニアの混合ガス、基板温度が例えば650°Cとする。つぎに、例えれば電子ビーム露光法を用いて、ポリシリコン膜を直径が例えば8nmまでの微細なポリSiドットに加工する。このポリSiドットが、微細分割型フローティングゲート42（電荷蓄積手段）である。その後、微細分割型フローティングゲート42を埋め込むかたちで、酸化膜44を、例えば9nmほどLPCVDにより成膜する。このLPCVDでは、原料ガスがDCSとN<sub>2</sub>Oの混合ガス、基板温度が例えば700°Cとする。この時、微細分割型フローティングゲート42は酸化膜44に埋め込まれる。平坦化が必要な場合は、新たに平坦化プロセス（例えはCMP等）を行うとよい。その後、ゲート電極8を成膜し、ゲート積層膜を一括してパターンニングする工程を経て、当該微細分割FG型メモリトランジスタを完成させる。

【0138】このようにSOI基板を用い、フローティングゲートが微細に分割されることについては、素子を

32

試作して特性を評価した結果、予想通りの良好な特性が得られることを確認した。

#### 【0139】変形例

以上述べてきた第1～第12実施形態において、種々の変形が可能である。

【0140】とくに図示しないがDINOR型、いわゆるHICR型と称されソース線を隣接する2つのソース領域で共有した分離ソース型のセルアレイから構成される微細NOR型セルなどの各種NOR型セルに対し本発明が適用できる。

【0141】本発明における“平面的に離散化された電荷蓄積手段”は、窒化膜バルクのキャリアトラップおよび酸化膜と窒化膜界面付近に形成されたキャリアトラップを含むことから、ゲート絶縁膜がNO (Nitride-Oxide) 膜なるMNOS型であっても本発明が適用できる。

【0142】本発明は、スタンダードアロン型の不揮発性メモリのほか、ロジック回路と同一基板上に集積化したエンベデッド型の不揮発性メモリに対しても適用可能である。なお、第12実施形態のようにSOI基板を用いることは、第1～第11実施形態のメモリトランジスタ構造に重複して適用可能である。

#### 【0143】

【発明の効果】本発明に係る不揮発性半導体記憶装置及びその駆動方法によれば、加速電荷が運動量（方向と大きさ）を維持しながら効率良く、しかも高速に電荷蓄積手段に注入される結果、高い書き込み速度を得ることができる。また、たとえばチャネル形成用と注入電界制御用として第1、第2ゲート電極を有することから、高い書き込み速度を得るためのゲート電圧設定が容易である。さらに、電子の引き抜きで消去することから信頼性および耐久性が高い、複数ビットの書き込みの適用が容易でありビットコストを低減しやすいなどの利点もある。

#### 【図面の簡単な説明】

【図1】第1実施形態に係る不揮発性半導体メモリのソース分離NOR型メモリセルアレイの概略構成を示す回路図である。

【図2】第1実施形態に係るNOR型メモリセルアレイの概略平面図である。

【図3】第1実施形態に係る図2のメモリセルアレイでA-A'線に沿った断面側から見た斜視図である。

【図4】第1実施形態に係るMONOS型メモリトランジスタの素子構造を示す断面図である。

【図5】第2実施形態に係る不揮発性半導体メモリの仮想接地NOR型メモリセルアレイの概略構成を示す回路図である。

【図6】第4実施形態に係るメモリトランジスタの素子構造例を示す断面図である。

【図7】第5実施形態に係るMONOSメモリトランジスタの素子構造を示す断面図である。

(18)

33

【図8】第6実施形態に係るMONOSメモリトランジスタの素子構造を示す断面図である。

【図9】第6実施形態に係るMONOSメモリトランジスタの製造において、LDD領域形成後の断面図である。

【図10】図9に続く、ゲート電極を分離し端面に絶縁層を形成した後の断面図である。

【図11】第7実施形態に係るMONOSメモリトランジスタの素子構造を示す断面図である。

【図12】第7実施形態に係るMONOSメモリトランジスタの製造において、ゲート絶縁膜の成膜後の断面図である。

【図13】図12に続く、片方のサイドウォール形導電層を除去した後の断面図である。

【図14】第8実施形態に係るMONOSメモリトランジスタの素子構造を示す断面図である。

【図15】第9実施形態に係るNOR型メモリセルアレイの構成を示す回路図である。

【図16】第9実施形態に係るNOR型メモリセルアレイの平面図である。

【図17】第9実施形態に係るNOR型メモリセルアレイにおいて、図16のB-B'線に沿った断面側から見た鳥瞰図である。

34

【図18】第10実施形態に係る微細NOR型セルアレイの概略構成を示す平面図である。

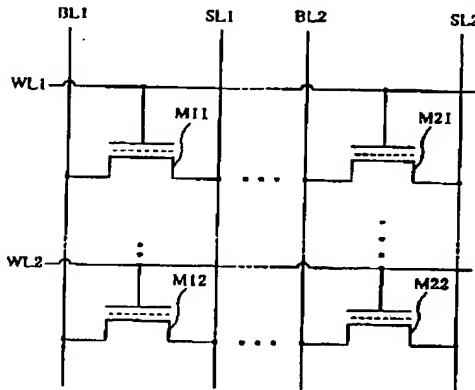
【図19】第11実施形態に係るSiナノ結晶型メモリトランジスタの素子構造を示す断面図である。

【図20】第12実施形態に係る微細分割FG型メモリトランジスタの素子構造を示す断面図である。

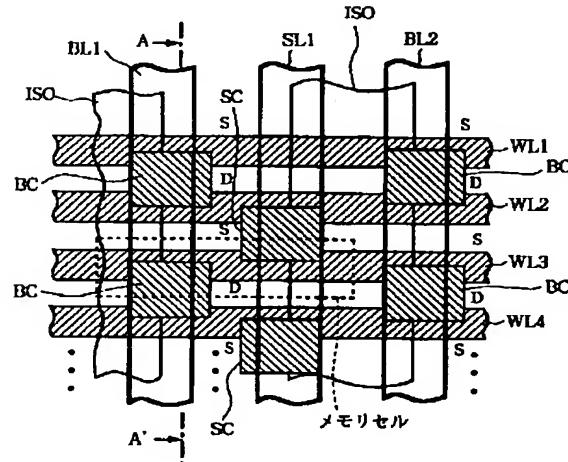
【符号の説明】

1…半導体基板またはウエル、1a…チャネル形成領域、2, S…ソース不純物領域（第2不純物領域）、4, D…ドレイン不純物領域（第1不純物領域）、6, 30, 40…ゲート絶縁膜、8…ゲート電極、10…ボトム絶縁膜、12…窒化膜、14…トップ絶縁膜、32…Siナノ結晶、34, 44…酸化膜、42…微細分割型フローティングゲート、46…半導体基板、48…分離酸化膜、50…シリコン層、60…nウエル、61…pウエル、ISO…素子分離絶縁層、PW…pウエル、M11～M22…メモリトランジスタ、S11, ST0等…選択トランジスタ、BL1等…ビット線、MBL1等…主ビット線、SBL…副ビット線、SL1等…ソース線、MSL…主ソース線、SSL1等…副ソース線、WL1等…ワード線、BC…ビットコンタクト、SC…ソースコンタクト。

【図1】

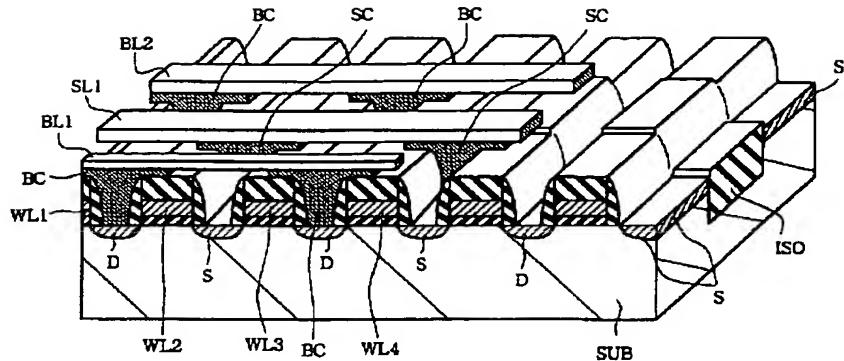


【図2】

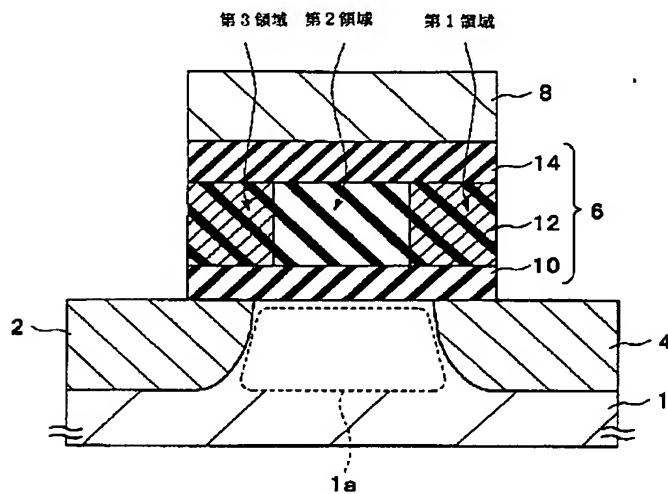


(19)

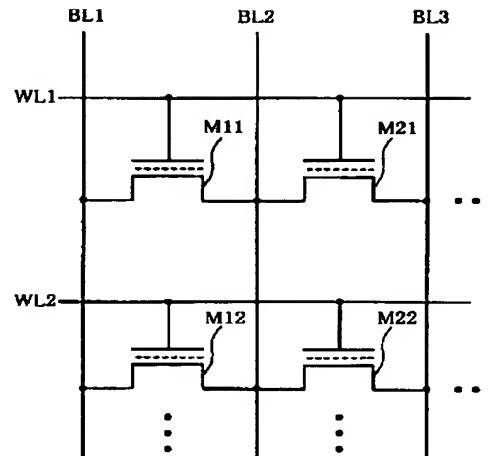
【図3】



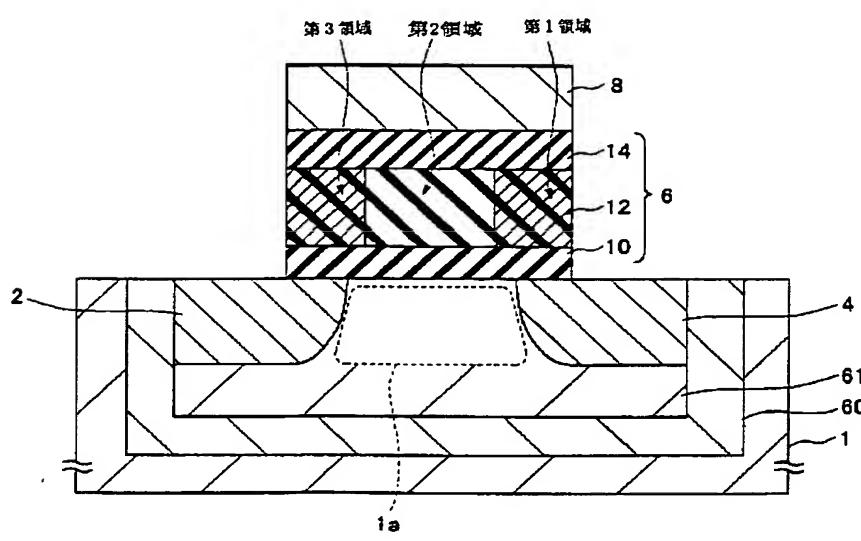
【図4】



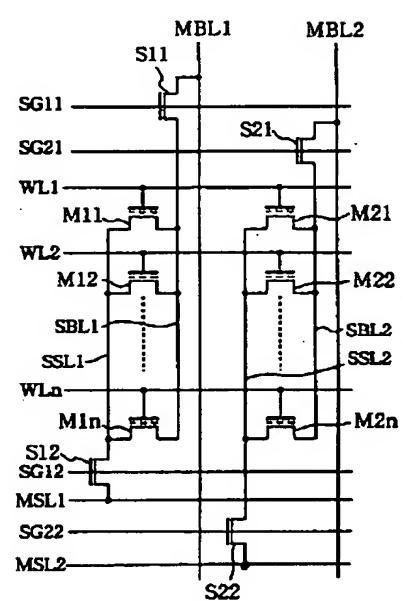
【図5】



【図6】

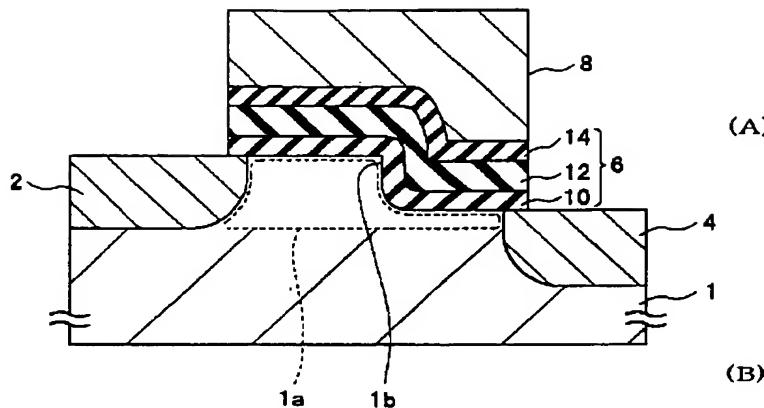


【図15】

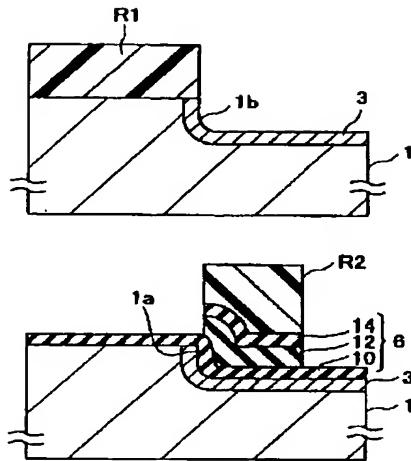


(20)

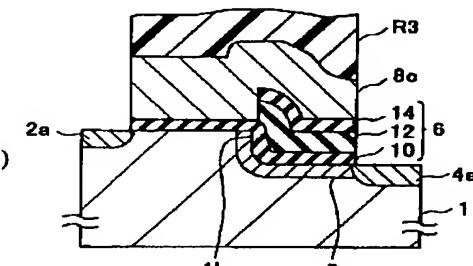
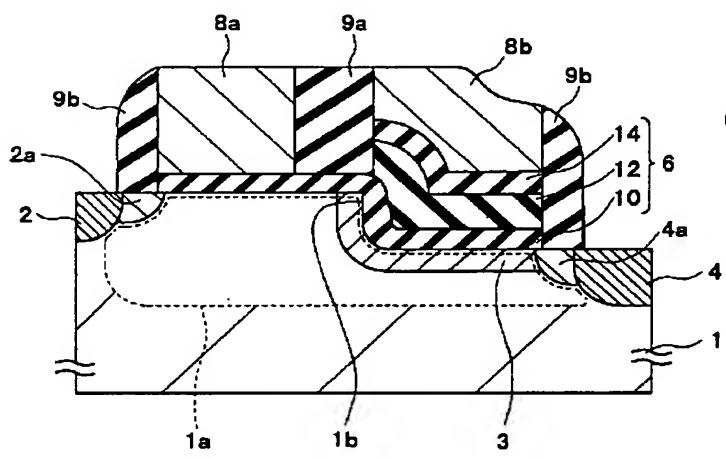
【図 7】



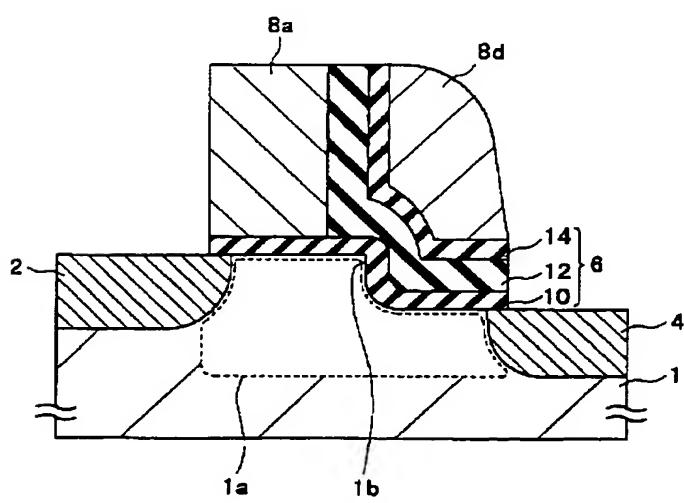
【図 9】



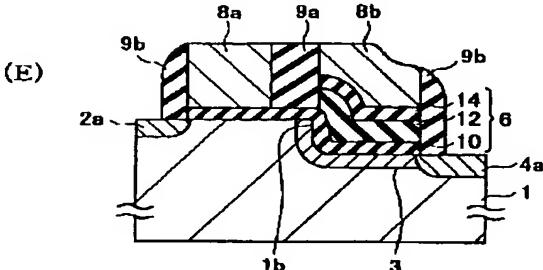
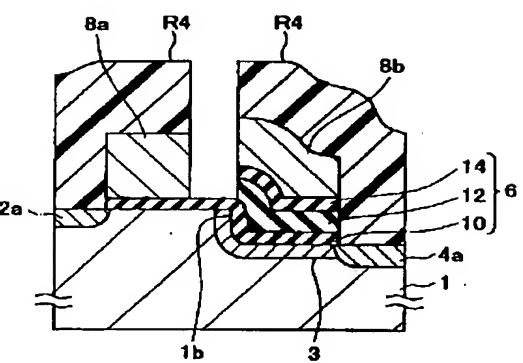
【図 8】



【図 11】

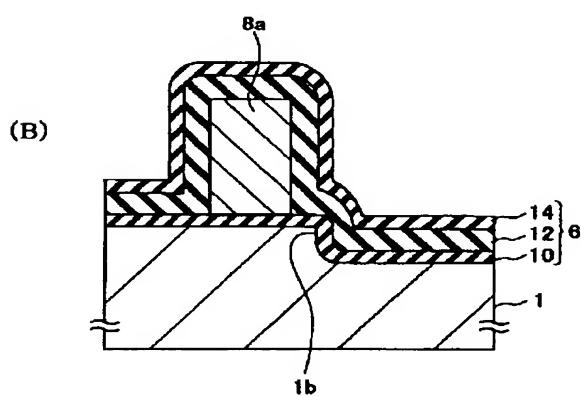
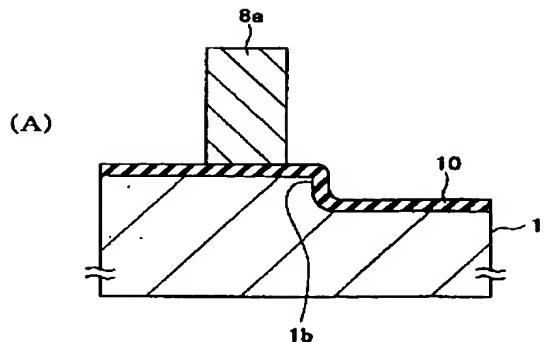


【図 10】

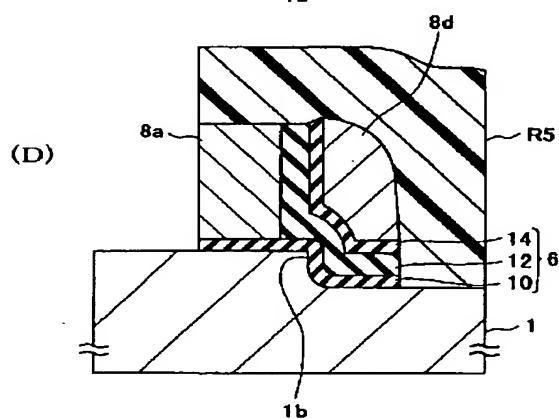
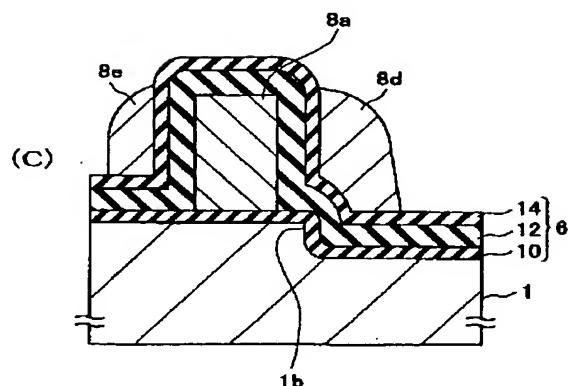


(21)

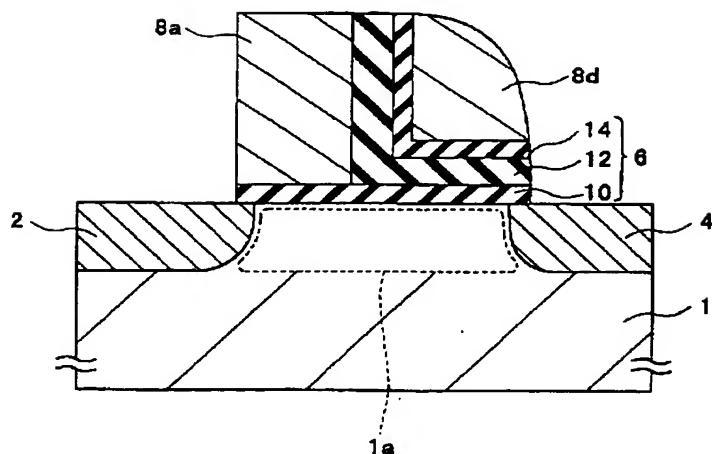
【図12】



【図13】

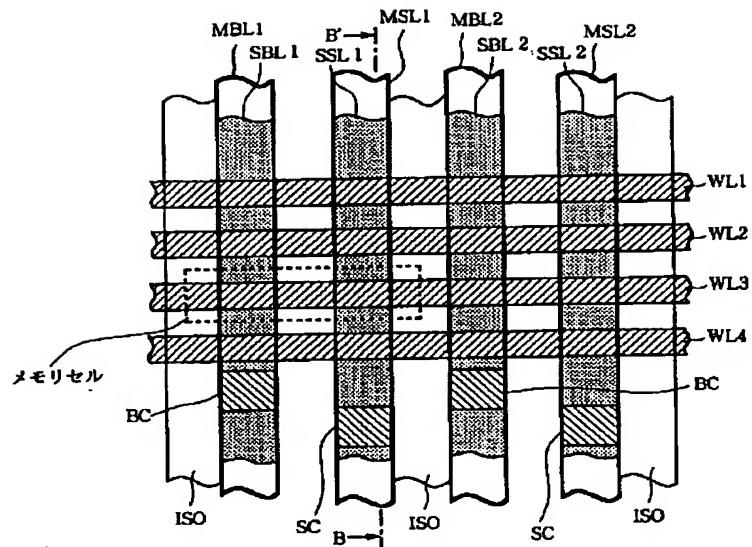


【図14】

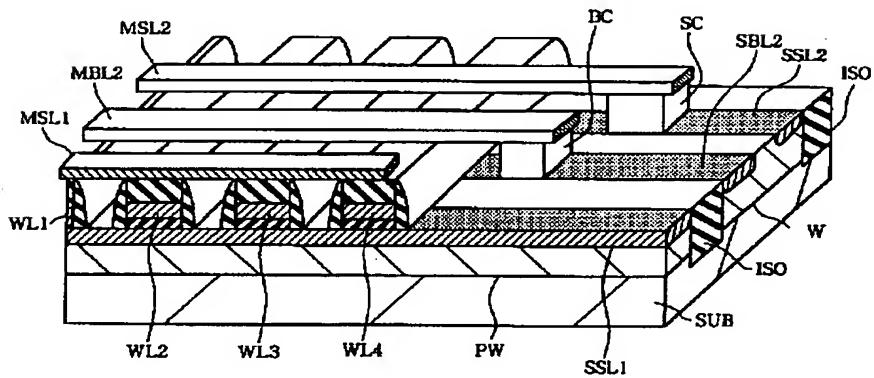


(22)

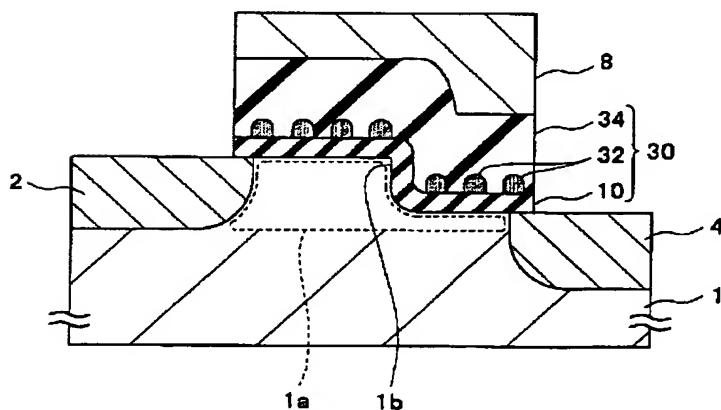
【図16】



【図17】

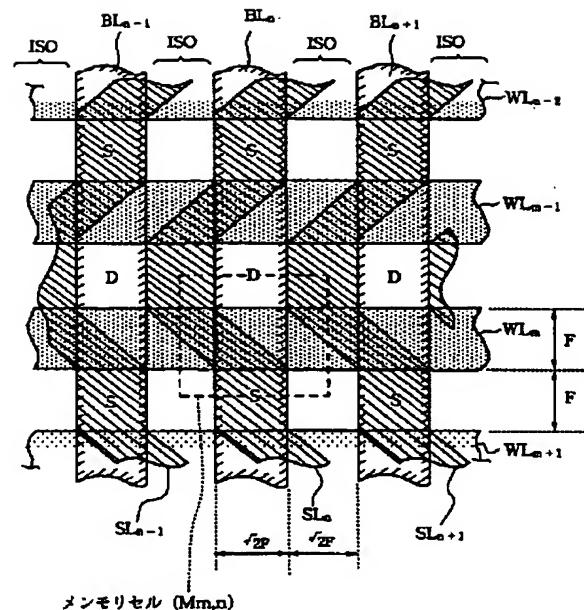


【図19】

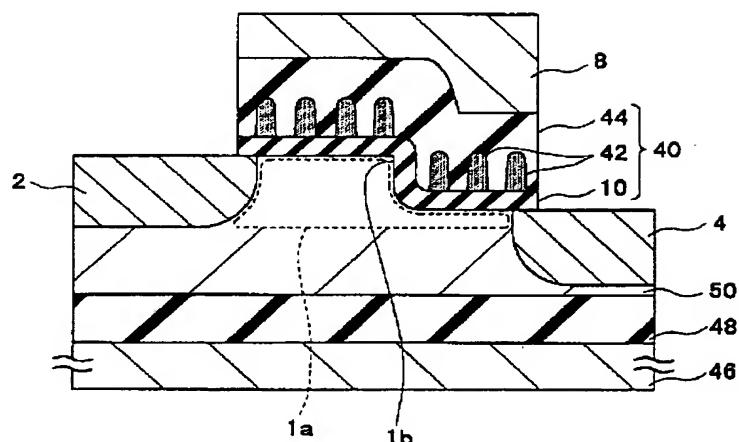


(23)

【図18】



【図20】



(24)

フロントページの続き

F ターム(参考) 5F001 AA14 AA19 AA34 AB02 AB03  
AC02 AC04 AC06 AC62 AD15  
AD17 AD18 AD21 AD23 AE02  
AE08 AF06 AF20 AG02 AG21  
AG30  
5F083 EP09 EP14 EP15 EP17 EP18  
EP22 EP49 EP55 EP63 EP68  
EP77 ER02 ER05 ER06 ER14  
ER19 ER22 ER23 ER30 GA30  
HA03 JA04 JA35 JA39 JA53  
KA06 KA12 MA02 MA06 MA20  
PR12 PR21 PR33 ZA21  
5F101 BA16 BA46 BA54 BB02 BB04  
BC02 BC07 BC11 BC13 BD05  
BD07 BD09 BD13 BD15 BE05  
BE07 BF02 BF05 BH02 BH03  
BH16